Liveness: an interactional account

Matthew Tobias Harris

Submitted in partial fulfilment of the requirements of the Degree of Doctor of Philosophy
Queen Mary University of London

Statement of originality

I, Matthew Tobias Harris, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below. I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material. I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis. I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

1st May 2017

Details of collaboration and publications:

Comedy Lab: human vs. robot was a collaboration with Kleomenis Katevas and Patrick GT Healey; individual contributions are noted where appropriate through section 4.3. Patrick GT Healey helped with the statistical analyses in sections 5.2.7 and 6.1.7.

Katevas, K., Healey, P. G. T. & Harris, M. T. (2014). Robot stand-up: engineering a comic performance. In *Proceedings of the Workshop on Humanoid Robots and Creativity at the 2014 IEEE-RAS International Conference on Humanoid Robots*

Katevas, K., Healey, P. G. T. & Harris, M. T. (2015). Robot comedy lab: experimenting with the social dynamics of live performance. *Frontiers in Psychology*, *6*(1253), 322

Liveness: an interactional account

Matthew Tobias Harris

Abstract

Live performances involve complex interactions between a large number of co-present people. Performance has been defined in terms of these performer–audience dynamics, but little is known about how they work. A series of live performance experiments investigate these dynamics, through teaching a humanoid robot some stagecraft, contrasting live and recorded performance, and spotlighting the audience. This requires the development of methods capable of capturing the fleeting responses of people within an audience and making sense of the resulting massed multi-modal data.

The results show that in live events interaction matters. Extending the idea that our experience of performance is shaped by interactions with others, namely by talking with people afterwards, analogous social patterns are identified within the event. Specifically, some of the interactional dynamics well established for close, dyadic encounters extend to performers and audience members, despite the somewhat anonymised nature of massed audiences. While individual performer–audience effects were identified, the primary axis of social interaction is shown to be between audience members. This emphasises how it is being in an audience – common across diverse performance genres – that shapes the experience of live events.

This work argues that the term liveness is ill-defined, but need not be. These interactional dynamics have a functional basis and depend solely on what is externally manifest. Understanding liveness in this way allows a perspicuous account – relating the perceptual environment within the event to the social contingency of experience – and can provide a systematic basis for design.

Contents

List of Figures

List of Tables

Acknowledgements

Coming off stage, I've often wondered to myself whether the audience would have had a better show if I'd pressed play on a particularly good studio take, and bobbed my head to my email inbox instead. My thing is a kind of improvised film, but anybody who has used a laptop on stage must have felt something similar at some point. My instinct is to make that laptop a better tool, or turn the interface into something legible for the audience. As a designer I'd need a way of reasoning about the situation... and not being able to reason about live events was the thread that kept on pulling.

The time and space to pull on that thread was made possible by the Research Councils UK's *Digital Economy* funding. With that, Queen Mary University of London established the *Media and Arts Technology* programme, gave me a scholarship to it, and I was finally able to reflect on ten years of practice to inform the next ten.

Most importantly, that thread got pulled on to a resolution. And that is in no small part thanks to Patrick Healey. First as co-director of that programme and then as supervisor, his inspiration and forbearance is there in every page of this work. I will not be the first to note the sheer intellectual rigour Pat brings, but perhaps distinct here is the shepherding of form. It would have been very easy for my PhD to become a portfolio of projects, as so much of my work before has been. I could have all too easily talked of the technological interventions I made, and talked up their significance. But I feel much more satisfied to be able to talk, for the first time in my life, of something closer to empirical fact than designer intent. Regardless, the research was a true collaboration and with the right kind of person that's a wonderful thing. Pat was the right kind of person. That was manifest in the experience had, and the results got.

Martin Welton, Graham White, Martin Barker and Andrew McPherson have all engaged deeply with this work and it's the better for it. As am I better for it. As second supervisor, yearly reviewer and final examiners respectively, the inter-disciplinary nature of this research warranted their differing perspectives, the challenges they made and, indeed, the validation upon submission and viva.

I took longer than I should have doing this. Looking back, it's pretty clear I've had a dysfunctional relationship with writing. But also looking back, some of my happiest memories are the support from Queen Mary's *Thinking Writing* group, in particular Kelly

Peake who has been there for me throughout. Long may their work continue.

Taking longer than I should has not been easy on those around me. Highlights of an engagement, two kinds of wedding, one child and then another do not make up for the sacrifices made by Marion, grandparents and others. It was my hubris that got them into this, and their selflessness that got me out of it.

So, thanks. So many thanks.

This work is supported by the Media and Arts Technology Programme, an RCUK Doctoral Training Centre in the Digital Economy; EP/G03723X/1

Being there 'live'

Music in concert, sport in a stadium, comedy in a club, drama in a theatre. These are live events. And there is something to *being there live*. Something that transcends the performance genre. Something that is about being there at an event, in the moment: caught in the din of the crowd, or suddenly aware you could hear a pin drop. This *something* is the topic of this dissertation.

The research works to transform the vernacular *live* into operational processes and phenomena. The goal is to produce an account of 'being there live' by describing how features common to live events shape the experiences had in them. This will necessarily pare away much of the richness of these genres of performance, as it will examples of individual events. What is sought is a foundational understanding, upon which such richness can be better considered.

The thesis is that an interactional analysis provides the most simple, clearly expressed and easily understood account of the liveness of live events. This is drawn from empirical investigation of the factors that contribute to the sense of being there live. In particular, it questions whether there are general patterns of interaction that can be used to generalise across live events.

Dissertation structure

First, the notion of the *liveness of live events* is set out. Chapter one uses existing literature to focus the research on human interaction. It does this by developing the examples of music in concert, comedy in a club and drama in a theatre and relating these to theoretical accounts of performance and liveness. Chapter one concludes with the need to establish a more perspicuous account of liveness than found in the literature.

The accounts of live events switch from written sources to direct observation in chapter two: a stand-up comedy event is described and analysed. The idea of performance as an interactional achievement is supported, most simply by the performer stating the act is going to be a dialogue with the audience, but more pervasively in the elaboration of some of the work identified in chapter one to this new data. The study of mass-interaction is limited however, and this chapter ultimately shows the need for a better understanding of

what to look for in human interaction, and a different approach to data collection.

The observational study of stand-up comedy demonstrated a gap in theory, method and instrumentation for the study of mass-interaction in live events. Chapter three addresses this gap by reviewing literature on human interaction. Applying the concerns of this literature in the context of live events leads to consideration of mass spectatorship. Is there a distinction between people who are merely massed together and *audiences*? This chapter argues that there is and that it consists in the specific kinds of social organisation involved.

The literature reviewed in chapter three also motivates an experimental programme. Chapter four presents the first, establishing Comedy Lab. A live performance experiment is staged that tests audience responses to a robot performer's gaze and gesture. This chapter provides the first direct evidence of individual performer–audience dynamics within an audience, and establishes the viability of live performance experiments.

The two main Comedy Lab experiments are presented in chapter five and six. Having successfully gained evidence of a social effect of co-presence in the first experiment, these two test the social effects of co-presence to the fullest extent practicable. This requires an expansion of the instrumentation, which opens chapter five. The basic premise of the experiment that follows is to have the performer as either an interacting party or not, and see what performer–audience and audience–audience dynamics are identifiable. The experiment contrasts live and recorded performance, directly addressing a topic that animates so much of the debate identified in chapter one. The data provide good evidence for social dynamics within the audience, but little evidence for performer–audience interaction. This emphasises that both conditions are live events, as even though the recorded condition is ostensibly not live, a live audience is present regardless and it is this that matters. Overall, the results affirm that events are socially structured situations with heterogeneous audiences.

The second main Comedy Lab experiment is presented in chapter six. The manipulation is now of the audience. The basic premise is to vary the exposure of individuals within the audience. The experiment contrasts being lit and being in the dark, when all around are lit or not. The data provide strong evidence for social dynamics within the audience, and limited evidence for performer–audience dynamics. Spotlighting individuals reduces their responses, while everyone being lit increases their responses: it is the effect of being picked out not being lit *per se* that matters. The results affirm that live events are social-spatial environments with heterogeneous audiences.

In pursuing Comedy Lab, challenges of capturing the behaviour of performers and audiences were repeatedly addressed. Beyond the issues of instrumentation already discussed, the data sources were diverse, and their combination and interpretation required original work throughout. Building on this work, a further contribution of method is made in chapter seven. A method to facilitate inductive analyses of performer–audience

dynamics is presented, along with the actual dataset visualiser tool developed. In the same way that video serves the study of face-to-face dialogue, augmented video and interactive visualisation can serve the study of live audiences.

The opening chapter set out the thesis that an interactional analysis should provide the simplest, most perspicuous account of the liveness of live events. In the chapters leading to the final chapter, eight, an empirical understanding of the interactional dynamics of particular live events has been put forward. This is now synthesised into an interactional account of liveness.

First, the Comedy Lab results are discussed as a response to the apparent paradox set up earlier in the dissertation. The programmatic hypothesis is that across live events, generalised patterns of mass interaction should be identifiable. However the interactional mechanisms that are well understood are dyadic and are found in everyday contexts. At first sight, live events – *massed! an escape from the everyday!* – would seem to be neither.

Following this, the interactional account of liveness is described. The concept of social topography is introduced and the nature of experience considered. It is argued that the experiments provide evidence that the kinds of experience-shaping conversations had after an event – "did you enjoy it?" – are happening, pervasively, during the event. With different interactional resources, they cannot be the complex verbal constructions of dialogue outright, but nonetheless they are there: moments of interaction that can change the whole trajectory of an experience. The interactional understanding of liveness put forward is then used to variously underpin, and undermine, some ideas of liveness encountered in the literature.

The exposition is completed with a consideration of how this account can provide a systematic basis for design. It argues that people have been long been alive to the issue of liveness and that technological interventions in particular can be powerful ways of reconfiguring experiences unique to live events. Further, as the dynamics of the interactions amongst audience members have been shown key to the experience of a live event, if practitioners attend to this directly new opportunities for intervention will open up.

Finally, the investigation of unfocussed interactions is discussed as future work, with specific challenges and risks informed by the Comedy Lab analysis. And it is noted that in measuring what is going between audience members, in making sense of those measures, in doing this with a much finer grain than anyone else has considered, and relating all this to experience... that this shows the need for a different orientation from performance studies, cognitive psychology, or even audience studies.

Chapter 1

On the liveness of live events

The adjective *live* used in the context of an event is a modern phenomenon. The first citation of the term in the Oxford English Dictionary is from the BBC Yearbook for 1934 as a response to the medium of radio. It was used as a qualifier for performance. It differentiated whether there were musicians in front of microphones in the studio, or a recording was playing instead. When performance was either people-in-sight or a gramophone-in-sight, which was being experienced was self-evident. But when all a listener had in the room was the sound from the radio's speaker, it wasn't self-evident. Such understanding fell to language: 'is it... *live*?'

While the term was established in opposition to technologies of reproduction, contemporary use does not bear this out. To a consumer of music, a 'live recording' – i.e., a recording of a live event – is a straightforward description rather than oxymoron. This new use is also a matter of status rather than what is qualitatively distinct about an experience. Yet the everyday pairing of *live* with *event*, of *live* with *performance* implies a distinction which is critical to this thesis.

A distinction is developed by considering experiences of live music, comedy and drama. This establishes the domain of interest: performance experienced in the context of an event and the role of interaction. Theoretical approaches are then reviewed, which leads to the conclusion that liveness is ill-defined, but need not be.

1.1 Live music, comedy and drama

If there is a sense of *being there live*, then the diversity of live events requires consideration. What might be common to the experience of live events? After all, musicians, comics and actors are doing quite different things.

1.1.1 Drama in a theatre

Drama at its best acts upon those present. In Carlson's (2003) autobiographical account –

> *I also have now and then experienced moments of such intensity that they might be called epiphanies... Such moments will be different for every theatregoer, but I feel certain that we all have them, and treasure them... [they] have served me as touchstones, as permanent reminders of what I have been seeking in a lifetime of theatregoing.*

The theatregoer interviews of Pasquier (2015) make clear such exceptional moments are just that, with audiences well versed in their own and others' disengagement –

> *When one's concentration goes, the body needs a release, by crossing one's legs, sitting up on one's chair... and coughing of course. That's the cacophony of failure. One senses the dispersion, people who start moving, changing position, who're leaning like this on their hand, who dip their head or look at others, you feel they're thinking 'shit, this is never going to end', who look at their watch, so it does show. I've got antennae...*

A common reference in theatre studies is Peter Brook's 1968 insistence that he could 'take any empty space and call it a bare stage', observing: 'A man walks across this empty space whilst someone else is watching him, and this is all that is needed for an act of theatre to be engaged' (quoted in Freshwater, 2009, p.1). Accounts that ground this concept of perception and co-presence with systematic analysis of actual performers or audiences are rare (though, see Gardair in section 1.2.2, *Observing interaction*). One workaround is to attend to the working language of performers, as done by Healey, Frauenberger, Oxley, Schober and Welton (2009). They claim performers routinely distinguish between "good" and "bad" audiences and between moments of intense engagement – "crackle", "movement", "lift" – and moments of "drop" and "drift". These moments define live performance in their view, but little is known about how, why or when they occur.

Probing what defines live performance motivated Reason (2004) to perform discourse analysis on theatre audience's post-performance pleasure-talk. His participants' experience of a theatrical production is shown to be shaped by the heterogeneity of the audience they were part of, and of the performers they watched. Perceptions of sameness and difference between those co-present in the performance space coloured much of their discussion. Reason notes that eye-to-eye and thigh-to-thigh contact is enforced by the venue featured in

the study; its small size made physical and mental awareness of their neighbours inevitable. What made these perceptions of sameness and difference particular to theatre was the relation between the base reality and the fictional world presented. For the performance attended, age was central to the drama, which related to differences in ages amongst actors and audience. The suggestion is that the particular pleasure and intensity of drama in a theatre is due in part to tensions within the audience and between the audience and the stage.

This systematic analysis of theatregoers' group talk gives weight to the personal introspection of theatre practitioner Tim Etchells. He writes "we're aware of the temporary and shifting bonds that link us both to the stage and to our fellow watchers, plugged into the group around and in front of us, the communal situation, sensing the laughter, attentiveness, tension or unease that grip us collectively, in waves and ripples, in jolts, jumps and uncertain spirals". Etchells also portrays limitations to this 'collective grip': "Sat watching, we spread-out, osmose, make connections. But at the same time, even as we do so, we feel our separateness, our difference from those around us, from those on-stage" (quoted in Freshwater 2009, p.7).

What also of Carlson's account: moments of such intensity they might as well be called epiphanies? With 'utopian performatives', Dolan (2006, p5) vividly develops these moments. She writes of "small but profound moments in which performance calls the attention of the audience in a way that lifts everyone slightly above the present, into a hopeful feeling of what the world might be like if every moment of our lives were as emotionally voluminous, generous, aesthetically striking, and intersubjectively intense". This reads as an idealised case of Etchell's communal situation, these individual experiences gaining from a shared, societal aspect. How they do that needs more development: could 'lifting everyone' apply to a distributed audience isolated from each other? The writing is compelling and there is much to consider. But, like Etchells's writing, it prompts many questions.

What are salient aspects to the experience of drama in a theatre? Answering this from the literature is surprisingly problematic. Theatre is generally said to have great affective impact on its audiences. However, accounts tend to be speculative and often result in exaggerated and unsubstantiated claims about theatre's influence and impact Freshwater, 2009, p.5, p.10. Nonetheless, the literature reviewed here suggests a social, embodied aspect is integral to the experience of theatre. The degree to which this is specific to theatre remains an open question. Much of this would seem applicable to events experienced in – using Reason's neat phrase – heightened social-spatial environments, such as a cinema.

Also an open question is what these words describing a communal situation might actually mean? Pasquier's interviewee has 'antennae'; Etchells says 'plugged in to the group' rather than simply 'part of'; what is it to 'osmose', or feel 'our separateness'; what is particular to performance in calling the attention of an audience? Pasquier's interviewee explains his

antennae well, listing actions of others that he picks up on and emphasises with. But those observed actions are only of disengagement. The suggestion of Etchells is that there is much more going on. Dolan suggests what is happening when theatre is at its best. There is a clear empirical challenge. What phenomena lie behind these descriptions, in the theatre, live?

1.1.2 Music in concert

An orchestral concert at its best has a sense of spectacle and occasion, according to audience researchers Pitts, Dobson, Gee and Spencer (2013). They continue: for audiences, there will be an expectation of quality; musicians are trusted to be committed and passionate about their work; promoters are trusted to provide a physically comfortable yet emotionally exhilarating experience. An alternative view is that classical musical concerts represent the archetypal boring experience in popular imagery. Here, non-connoisseurs feel distant from the experience being offered, according to experience scholars Carù and Cova (2006). This self-report from a participant in their study has a clear sense of the highs and lows possible, even within the same concert –

> *"It's beautiful! I really like the music, it's truly magical. I don't know what I'm thinking because I can't control my thoughts... I can't even feel my backache anymore. Physically I feel very relaxed. I have ABSOLUTELY no interest anymore in the place or the people."*
> ...
> *"I've been thinking about work and losing concentration. The experience has been ruined."*

The epitome of the experience of live listening is a sense that highly accomplished players are within reach, yet still at a professional distance, according to Pitts et al. (2013). This sense is linked to a view that all participants are connected during the performance, but take their separate responses away at the end of the concert. Such connection somehow works at the level of groups. Players and audience are two distinct groups; individual player–audience connections are the exception as the event is structured to separate their experiences. Particularly relevant to a generalised sense of being there live, the listening experience is asserted to be strengthened by the collective act of engagement, although how this is manifest is not addressed. This high-level analysis is based on extensive consultation with players and listeners to gain an understanding of how orchestral concerts are experienced. Closer to the underlying data, they note the importance to many audience members of watching the body language and expression of those on stage. This aligns with experimental findings that demonstrate that people depend primarily on visual information when making judgments about music performance (Tsay, 2013).

Musicians' thoughts and perceptions during performance of classical music were investigated by T. Clark, Lisboa and Williamon (2014). Players reported speaking of heightened

concentration and sense of awareness of other players while on stage, and generally characteristics associated 'with the experience of the state of flow'. Less successful performances had players speaking of inappropriate, scattered or negative focus. Fixating on this negativity would then further hinder their ability to perform. Audiences impacted on performance quality both positively and negatively, with 'communicating with the audience' being an appropriate focus for some musicians during performance; smaller than expected audiences disheartening some performers; moving and talking in an audience mid-performance negatively affecting some musicians. The performer knowing members of the audience could have a strong effect, although this is likely heightened in the critical, conservatoire environment their participants were drawn from. A contrasting view is that during the actual playing, the performer's sense of the listeners is largely suppressed. 'Except, of course, when they misbehave', in the view of Rosen (2002, p.122). Drawing on his career as a concert pianist, he goes further: 'a slightly scandalous effect of a breach of decorum' is suggested if the performer 'betrays' an awareness of the audience (p121). On the audience misbehaving, for Rosen the 'audience breathes silently in darkness' (p.121), 'a cough is the basic sign of inattention' (p.127), and 'only when the piece is over may the public manifest its existence and express a collective and noisy opinion' (p.121).

The live music event has been conceptualised primarily by authors focused either on music or society. The results tend towards taxonomies of musical genre or cultural theory. This shifts the focus from actual experiences. An account that attends to the event itself is Webster's (2011) exploration of the practices and experiences of promoters in the UK. Drawing on her own and others' ethnographic work, she presents a typology grounded in the idea that a live music event is a social experience that involves simultaneous production and consumption. While presented as four categories, it is easier understood as a continuum between presentational and participatory forms (c.f. section 1.1.4, *Liveness?*). Webster goes on to separate factors affecting audiences enjoyment of concert events into 'anticipated behaviour' and 'actual behaviour'. Anticipated behaviour would depend on an individual's previous experience of live music events. Actual behaviour would be based on 'background modifiers' such as seating or standing and other factors of the spatial environment, and 'dynamic modifiers' such as the behaviour of other audience members. (p.155)

Getting at actual, lived experience during a concert appears problematic. There is a paucity of work focussed on understanding in-situ experiences of concerts. Pitts et al. (2013) provides a case in point. The central question is about experience. Answers are drawn from an extensive evidence base. Nonetheless the analysis concentrates on the experience surrounding, rather than in, the concerts – in essence, what is useful to the arts institution in terms of marketing and audience development.

1.1.3 Comedy in a club

Life-affirming and absurd. The attraction of live comedy is evoked in these two extracts from participants in Lockyer and Myers's (2011) investigation of live stand-up comedy from the audience's perspective –

> *I like a really good laugh and I like to see somebody brave it, as it were, in front of an audience. I think there's nothing more life affirming than being with a group of other people watching somebody entertain you on a stage.*
>
> *It's that collective laughter and that sense of being raucous where everyone's yelling, in hysterics at this one person, and the one person knowing what they're doing is great and obviously working the room. It's that collective sense of belly laughing and the absurdity of it all.*

There are broadly two kinds of live comedy in a club, stand-up or sketch. Stand-up comedy is typically configured as a single person standing-up in front of the audience, addressing them directly. Sketch comedy is typically configured with performers operating as a group, supported by props or other scenographic elements, with direct address to the audience the exception. This section focuses on stand-up comedy as it is the more distinct form of live event. Sketch comedy shares many features with drama.

There are five salient aspects to the experience of live stand-up, according to Lockyer and Myers's (2011) thematic analysis of semi-structured interviews they conducted with eleven stand-up comedy goers. Interaction features heavily, defining at least two.

Opportunities for interaction: interaction between audience members appeals to live stand-up comedy audiences; interaction with accompanying friends, family, and partners was encouraged and interpreted as a positive addition to the experience. Stand-up comedy events are often configured for intra-audience interaction. Seating may be around tables with social drinking encouraged. The performance format may involve many short performances and so frequent breaks. In the words of one respondent talking about a particular venue: "you're around the table, you've got a few drinks, you've had a good chat with your friends, and then a comedian comes on and they interact with the audience, and then they'll go off and you've got another few minutes and then the next. So it's a whole different night out because it's more communicating with your friends rather than, like going to the cinema where you're sitting in a dark room in a straight line". In contrast, interaction between the stand-up comedian and the audience was generally discouraged, and in some cases negatively perceived. The respondents were conscious that participation fostered upon audience individuals – being 'picked on' – was done for the entertainment of the rest of the audience. Part of the unpredictability experienced, and desired, in live stand-up comedy also stems from such audience participation, or heckling from the audience to the performer. The respondents were conscious of performer and audiences engaging with the issue of

Attends With	Response %
Friends	68
Partner/Wife/Husband	42
Family Members	21
No One – attends on own	4
Work Colleagues	2
Carer	1
Other Comedians	1
Housemates	1
Acquaintances	1

Table 1.1: *Who do you usually go to see live stand-up comedy with? 130 respondents; multiple answers were permitted.*

participation. This was seen in some choosing 'safe' seats and some ascribing performers a skill in identifying audience members likely to 'play along'.

Sharing the comic experience: experiencing comedy with others, as it happens, is sought and valued. Laughing was said to have a different dynamic when with others than when watching a comic alone, ie. on television. Laughing together was said to have value in establishing and maintaining friendships. The "collective laughter...hysterics at one person...working the room" quote above expresses this well. In another respondent's words: "A lot of it, I think, is in that moment. A lot of it is mass audience, a lot of it's being with other people that are all laughing at the same time. I think you get an awful lot out of that. So I think a lot of it is being caught up in the moment with other people that are also finding something that you find funny".

The environment, already mentioned in terms of seating around tables vs. straight rows of seats, warrants *proximity and intimacy*. The spatial arrangement of the stand-up comedian and the audience is significant. Respondents expressed that they enjoy being close to the stand-up comedian, attributing this to increased senses of being part of the performance, of rapport with the performer, or of overall intimacy in the performance space. This shares a sense of connection with the performer, seen in the penultimate aspect of *respect for the comedian*. Appreciation of the live stand-up comedian goes beyond appreciation of their performance skills to a particular respect. The lone stand-up comedian is respected for the courage to stand before an audience which is expecting to be amused, and then entertain them single-handed. "it's a phenomenal talent to hold that many people's attention for that long". The last aspect is *expecting the unexpected*, where the potential for live stand-up to have unexpected and unpredictable moments is distinct and appealing.

Attendance of stand-up comedy events was regarded by many of these eleven interviewees as a social event. The authors also have a corpus of online survey responses drawn from 248 respondents. One question from this gives further insight of the sociability of attendance. Asked 'who do you usually go to see live stand-up comedy with?', just four

Venue	Response %
Small Arenas/Theatres	55
Small Comedy Clubs	51
Medium-sized Comedy Clubs	46
Large Arenas	38
Small Rooms in Pubs	31
Other	4

Table 1.2: *Where do you usually see live stand-up comedy? 121 respondents; multiple answers were permitted*

percent replied they would ever go alone. Table 1.1 shows all responses.

Venues where stand-up comedy is performed were also categorised by Lockyer and Myers (2011), who then gathered responses of attendance and preference via online survey. Table 1.2 lists the venue categorisation and responses to the question "Where do you usually see live stand-up comedy?". Notably, preference responses did not align with the responses of actual attendance shown here. Small venues are preferred; preference percentages were 15% for large arenas and 39% for medium-sized comedy clubs.

The live stand-up event is typically conceptualised through the motives and techniques of comedy performers. These in turn draw upon theories of comedy which relate the efficacy of a joke to its delivery. The setting and context in which the joke is told, the competency of its delivery, the identity of the teller, and the audience to which it is told all factor, in the account of Palmer (1994). This can be seen developed in Double's (2005) definition of stand-up comedy, part of which states that 'it happens in the present tense, in the here and now. It acknowledges the performance situation. The stand-up is duty-bound to incorporate events in the venue into the act. Failure to respond to a heckler, a dropped glass or the ringing of a mobile phone is a sign of weakness which will result in the audience losing faith in the performer's ability'. Accounts of live comedy events develop the importance of the audience beyond an arbiter of a joke's success. A relationship between comedian and audience is proposed as central to the event. Double (2005), drawing in part on his own experience as a comic, suggests 'take the audience away from stand-up comedy and it starts to look weird ... stand-up comedy without an audience is only half there' (p.106). He defines stand-up comedy as involving direct communication between performer and audience, characterising an intense relationship, "with energy flowing back and forth between stage and auditorium". It's a kind of conversation made up of jokes, laughter and "sometimes less pleasant responses" (p.19). Rutter (1997), using conversational analysis techniques, exposes systematic efforts by compères in a club he studied to establish a relationship between audience and the forthcoming performer. Scarpetta and Spagnolli (2009) analysed the same comics performing in front of different groups and concluded an interactional context was created through their performance, allowing the comics to tailor

their delivery and encourage the audience's appreciation and acceptability of the material.

1.1.4 Liveness?

Common to these forms of events seems to be an intensified experience. This is written about sometimes as individual subjective response, and sometimes as a group phenomenon. The Carù and Cova participant (section 1.1.2, *Music in concert*, "It's beautiful! ... the experience has been ruined") is a great example of individual subjective response, with extreme highs and lows entirely vested in her self-reported internal processes. As group phenomena, Pitts et al. talk of a collective act of engagement (section 1.1.2), the participants of Lockyer and Myers emphasise the experience of collective laughter (section 1.1.3, *Comedy in a club*), Etchells a 'collective grip' (section 1.1.1, *Drama in a theatre*), Dolan has individual experiences gaining a shared, societal aspect (section 1.1.1), and so on. Even in the self-report that exemplified individual subjective response, there is an individual vs. collective aspect: "no interest anymore in the place or the people".

Having a sense of others in the event is a common theme. This can be extreme, exemplified by the Lockyer and Myers's (2011) participant – "that sense of being raucous where everyone's yelling, in hysterics at this one person, and the one person knowing what they're doing is great and obviously working the room". There is much to analyse in one person having an awareness of the performer 'working' another person in the audience. The sense of others may be suppressed, as observed by the concert pianist Rosen (section 1.1.2, *Music in concert*). It is not clear whether this suppression is an act in itself, requiring effort. The suppression cannot be complete however as the audience 'misbehaving' is picked up by the performer. Also how the sense of others becomes manifest may be surprising, for example the importance of the visual in judging performance of an art form whose medium is sound (section 1.1.2, *Music in concert*).

Generalisation is difficult. Consider how the context of watching performance changes across these examples. The particularities of that context change with each kind of performance. For instance, the convention of stand-up comedy is to acknowledge the performance situation, incorporating events in the venue into the act. This would be a marked act in most theatres, however. Convention there keeps the fictional world on-stage separate from the auditorium. This suggests direct address to, from, or between performers and audience is not fundamental to a sense of being there live. The particularities of that context changes with each kind of venue, too. Live events are experienced in small clubs and expansive stadiums, with audience sizes orders of magnitudes apart. Experiences of live music and comedy in particular have different characteristics across these scales (Webster, 2011; Lockyer & Myers, 2011). This suggests any particular configuration of performer and audience is not fundamental to a sense of being there live.

Broader characterisations can be made. Webster's typology (section 1.1.2, *Music in concert*)

encompassed the heterogeneity of music events and can also be applied beyond music. Drama in a theatre would fit the *presentational* type[1]. Comedy in a club would fit the *participatory presentational* type[2]As this typology is grounded in the idea that a live music event is a social experience that involves simultaneous production and consumption, and having a sense of others is established above as a common theme, this suggests all live events are social experiences.

Putting this together, the programmatic hypothesis can be made. There is something qualitatively distinct to the experience of live events, and this *liveness* depends on the social effects of co-presence.

A social hypothesis contrasts with influential accounts of live performance. Bennett's (1997) model of the audience's experience of theatre focusses on socio-historic-economic specificities – "the cultural conditions that make theatre and an audience member's experience of it possible" (p.vii). This moves the analysis away from individuals and the actual goings on during an event. Any social aspect during the performance is sidelined; ultimately, it is a model of individual subjective response. Auslander (2008) allows communality within an audience, but denies it has any relation to the performance: the quality of the experience of community derives solely from the specific audience situation. The analysis is cultural-economic, generalising the performance as spectacle and devaluing any communal bond as "likely to be little more than the common consumption of a particular performance commodity" (p.64). This abstracts away from a consideration of what is actually going on in-situ. What if the live performance could shape that specific audience situation? Mutual influence and spontaneity is similarly deemphasised. The argument is that productions cannot afford to vary much in response to different audiences. Even in performances that solicit input directly from the audience, valuing that 'live' influence may be misplaced: the performers may use the suggestions as pretexts for doing something they wish to do anyway; one can enjoy performers engaging with audience suggestions while being outside of that audience itself, such as watching on television (pp.68-72). In this account, any social aspect during the performance is minimised to irrelevance; ultimately the argument denies any qualitatively distinct sense of being there live.

The heterogeneity of live events leads to contrasting analyses. Making a similar argument to the earlier point on differing venue sizes, Auslander (2008, pp.66-68) reasons that the value of presence cannot be equated with a particular relationship between audience and performer. Further, being present is not universally better, if for example a football game can be seen and the details followed much better on television than at the stadium.

[1] "Whereby the attendees are consumers with little or no contribution to the production of the music or the spectacle on stage; a symphony concert, for example. Akin to the Western classical concert tradition, the attendees will usually remain silent throughout the duration of the music performance. Webster, 2011, p.154

[2] "Whereby the consumers and producers influence each other in a continuous feedback loop but where the producer is still the focus. There will generally be a greater level of audience participation, both physically and verbally. Performers often highlight the participatory sections through actions such as holding a microphone out to the audience or initiating call and response patterns. A rock gig, for example. Webster, 2011, p.154

In such arguments, diversity is used to dismiss aspects of performance. In other arguments, diversity gives broader insight into how performance works. Bennett (1997) arrives at her model of the audience's experience of theatre in this way. However the focus on cultural particularities leaves little that is generalisable for a sense of being there live. What is needed is something like the approach of Fischer-Lichte (2008), who draws upon the diversity of events to create a concept of performance that starts with the encounter of individuals. It is a from-first-principles account of the experience of performance.

Overall, a coherent account of liveness cannot be deduced. Consistent with the illustration in the introduction of being caught in the din of the crowd or suddenly aware you could hear a pin drop, intensified experience and awareness of others appear to be critical. However there is neither a satisfactory account behind this nor the evidence base to argue convincingly from. Accounts that emphasise co-present sociality show promise in their engagement with the in-situ conditions of a live event. Accounts based on a cultural analysis seem less useful, by definition less generalisable and abstracted from these in-situ conditions.

Overall, there is a paucity of literature discussing what is actually going on at an event, during the performance (though see Reason and Sedgman 2015). Freshwater (2009, p.11) suggests this may be due to the difficulties of doing justice to the fleeting phenomenon of collective and individual audience response. Indeed, any claim of a direct, empirical relationship to actual, lived experience is problematic. Reason (2010) suggests the experience of a theatrical performance, if thought of in terms of the primary reactions and individual responses, is essentially unavailable to the knowledge of others – and perhaps even to our own reflective or conscious knowledge of ourselves.

To work towards a perspicuous account of liveness, the speculative theory needs to be challenged with empirical enquiry. The next section sets this up, starting with a detailed consideration of Fischer-Lichte's (2008) concept of performance.

1.2 Theorising liveness

1.2.1 Speculative interaction

Is there something qualitatively distinct to the experience of live events, and is it rooted in the social effects of co-presence? Fischer-Lichte's concept of performance suggests so. It consists of four arguments.

> *"First argument: A performance takes place in and through the bodily co-presence of actors and spectators. For in order to bring it about two groups of people who act as 'doers' and 'onlookers' have to assemble at a certain time and at a certain place in order to share a situation, a span of lifetime. A performance comes into being out of their encounter – out*

of their interaction."
— Fischer-Lichte (2010)

The social effects of co-presence explicitly drive this argument. Reactions to the performance are in part externally manifest, and the perception of such responses results in perceptible reactions of others. The exposition develops the position that 'performance' itself is a result of feedback processes and is bound to their continuing operation. In essence, preconditions are set, and then performance arises through interactional dynamics.

The reactions cited in the exposition are decidedly overt – "giggling, laughing, shouting, yawning, snoring, sobbing, crying, eating, drinking, commenting on what is happening, getting up, running out, slamming the doors and so on and so forth". Fischer-Lichte's basic premise however is that all forms of physical encounter between people stimulate interactions, "even if their shape is not always plainly evident" Fischer-Lichte, 2008, p.43. By adapting Watzlawick, Jackson and Bavelas's first axiom of communication that "one cannot not communicate" into "you cannot not react to each other", she touches upon social science detailing human interaction. In such views, interaction is pervasive and fine-grained, often requiring 'micro' analysis to identify the cues of interaction (e.g. Bavelas, Black, Lemery and Mullett 1986). An intuitive example is gaze in an auditorium, where the eyes or head position may provide a 'reading' to others of what is being attended to (e.g. Frischen, Bayliss and Tipper 2007).

If performance is a social process, an analysis of liveness has to question how individual experience is shaped by interactions with people, likely within a mass of people. The mediating conditions of the live event are critical to these interactions, and so an important part of the analysis.

"Second argument: The materiality of a performance, its spatiality, corporeality and sound quality is brought forth by and in the course of the performance. Hence follows the paradox of performance. It is ephemeral and transitory. However, what appears and takes shape in its course, comes into being hic et nunc *and is experienced as present in a particularly intense way."*
— Fischer-Lichte (2010)

The interactional dynamics established in the first argument are now used to create a 'performance space'. The performance space is not, then, a fixed property of what is there. To Fischer-Lichte it is the possibilities opened up between those present, unstable and changing; it is what matters.

Overall, this argument can be read as the creation of place – as in, we are located in space but act in place (Harrison & Dourish, 1996). Such consideration of how action transforms the pre-existing conditions aligns with sociolinguistic views where meaning and context are something achieved through interaction – such as the Conversation Analytic approach to

the construction of institutions (Heritage, 2005). To this end the shared construct of place built between the participants is used to organise and structure movement and perception. Meaning is achieved from how these resources are 'realised, evaded or counteracted'. This meaning-making applies in part to the 'materialities brought forth' as outcome of this second argument, with the third argument based entirely upon it.

How the performance (space) is 'experienced as present in a particularly intense way' involves atmospheres that 'pour into the space'; ecstasies of objects that 'demand attention'; phenomenal bodies that in their 'different physiological, affective, energetic and motorial states' are able to evoke these states in others; and semiotic bodies that are 'indissolubly bound' to the phenomenal. Of these, only the discussion around evoking state in others has a co-present aspect, and the mechanism is unclear (phenomenal possibly suggesting automatic rather than social processes, but this is speculation).

If what matters within the live event is an interactional achievement, an analysis of liveness should be able to track the development (and dissolution) through the event. An analysis of liveness must question whether this active construction shapes the experience of the individuals involved. If there is something qualitatively distinct to live events, could this be why?

> *"Third argument: A performance does not transmit pre-given meanings. Rather, it is the performance which brings forth the meanings that come into being during its course."*
> *— Fischer-Lichte (2010)*

This argument is predicated on the conditions of every performance being different, most plainly in that the spectators change (and if not the actual spectators themselves, then their circumstances and arrangement). Following the earlier arguments, this necessarily leads to different outcomes. Again, there is alignment with social science views where context is actively re-created on each new occasion (Heritage, 1984, p.106-9).

Fischer-Lichte focusses on perception and speculative phenomenology. Spectators oscillate between perceiving 'presence' and 'representation'. This oscillation draws the attention of the perceiving subject to the process of perception. As a result what will be perceived and the meanings produced become hard to predict. Nonetheless, it is hard to see how this is different from an individual interpretation of a text. It doesn't directly involve the dynamics between those present.

This in many ways is a restatement of the first two arguments, used to differentiate this concept from a tradition where the purpose of performance is to convey specific pre-given meanings. The experiential aspect is considered separately, as part of the next argument.

> *"Fourth argument: Performances are characterised by their eventness. The specific mode of experience they allow for is a particular form of liminal experience."*
> *— Fischer-Lichte (2010)*

This argument is in part about the conditions for interaction, what those conditions can facilitate, and what, if anything, can result afterwards. Due to interaction, each person present co-determines the course of the performance. Similarly, no participant will have complete control of it, in particular the spectators. This is the starting point for it being "event" – it happens to you – and "liminal experience" – you experience it as both happening to you and you happening in it.

A further kind of liminality comes from the oscillation of perception that forms part of the third argument. The shifts in perception draw attention both to the moment of transition and the different modes of perception. The repeated shifts in perception lead to 'a strange collapsing of oppositions' between these two modes. Liminality is an important concept in performance studies, traceable to ethnographic study of ritual that emphasised symbolically loaded threshold experiences and rites of passage Fischer-Lichte, 2014, p.42. This argument emphasises liminality as a state of exception as much as something between two states.

More broadly, this argument is about establishing the primacy of the event as organising concept and what modes of experience this facilitates. Fischer-Lichte argues for a privileged category of experience opened up by the dynamic processes of performance. The core of this is everyday human interaction: that interaction is intrinsic to humans and we make our world intelligible through our ongoing, mutual actions within it. This doesn't stop when performance starts, rather performance is built upon it.

These arguments provide a theoretical account of live events consistent with the commonalities and notion of liveness identified in section 1.1, *Live music, comedy and drama*. In this account, interactional dynamics between those present are unavoidable and transform the experience for all. Unfortunately, the discussion of interactional dynamics is largely speculative. Evidence is needed.

1.2.2 Observing interaction

Street performers have little resource beyond their bodily co-presence with a crowd. This makes street performance events an ideal place to look for evidence of the role of interaction in a live experience. Gardair (2013) did this and details many interactional cues that constitute performance as interactional achievement. Key to her analysis was asking what makes people act as audience members. What is an audience? How do you make one? What does it get you?

The choice of street performance gives a particular power to the study. In the site studied, the performers are working in an environment that is not configured for performance. They advertise on the spot to passers-by, who are there for other reasons. This allows the study of performance when stripped of the structures that frame it as performance. Pragmatic measures of success also distinguish street performance; in the

performers' ability to establish the performance itself, and in the financial success – or lack of – that ensues.

The study required an empirical understanding of what performers and audiences did during the event. Field notes were recorded and interviews with participants were conducted. However these explanations often proved unreliable upon systematic review of actual behaviour during the performance – the use of video as reliable witness was critical. Note the recurring problem of the availability of an experience to others (e.g. 1.1.4, *Liveness?*). Further, video recording allowed consideration of procedural detail within and across performances. Gardair's systematic review approached the methods of ethnomethodology, drawing in particular upon the ways the disciplines surrounding Conversation Analysis study embodied interactions and make sense of social practices and public encounters (e.g. Heath and Luff 2010; Heath and vom Lehn 2004; Livingston 1987.

The basic premise is that only by considering the performance events as a specific kind of social practice could the observed phenomena be explained. Street performers are shown to create their performance space, build their audience, and elicit payment. Audiences are shown to collectively produce appropriate responses at the appropriate moment. Analyses of performance as the enactment of a text, or an audience as an assembly of spectators, are argued to be inadequate to explain the observed behaviour. An analysis adequate to the observed behaviour was based on principles of social encounters, such as how people often group themselves into set patterns (Kendon, 1990) and that people design their behaviour for others (Goffman, 1959). In this analysis, performances are opportunistically established as 'frames' of participation managed by interaction. Those within maintain the framing and adhere to its rules. Those outside recognise it taking place and avoid disturbing it.

The analysis is made through more than fifty excerpts illustrating the patterns of interaction observed. To address getting people's attention, for example, the argument has to resolve how street performers make their actions recognisable as the pre-cursors to a performance. For this, efficacy is judged using, amongst others, the concept of body torque (Schegloff, 1998) as manifest in passers-by. This allows a determination of engagement, where some – Gardair's *lookers* – only change their upper body segments while others – Gardair's *watchers* – re-orient their entire body. What matters to engage passers-by is shown to be how the actions are designed for them and how their responses are integrated into the act. An ostensibly testing-the-mic noise changes to a 'hello' precisely at the point a head of a passer-by turns towards the noise. The presence of a chainsaw only has the effect of moving people when the performer checks who is looking at it and orients a gesture to them. It is these specific kinds of interaction that makes the actions and objects performative, rather than merely anomalous.

The key insight is that the performers spend the majority of their time transforming passers-by into an audience, and this work is rewarded in the successful solicitation of

monies at the act's end. Gardair's claim is that an audience is an ongoing interactional achievement built up through a structured sequence of interactions between performers, passers-by and audience members. Through these interactions performers get people's attention, achieve the recognition that what is going on is a performance, build a collective sense of audience membership, establish moral obligations to each other and the performer, and train the audience how to respond. Gardair argues that the self-identification of spectators as audience members is key; as groups are formed, individual behaviours alter. So while outwardly similar a crowd and an audience behave in different ways. Street performers know this and exploit it, working not to merely assemble spectators but to engineer the social conditions that will maximise the impact of their act and, ultimately, financial reward. These social conditions make collective responses possible and lead to a 'ripple effect' where the first donations at an act's end engender donations from those who would not have otherwise.

In demonstrating the pervasive relevance of interaction to the achievement of a successful performance, the findings of the study often reify what was speculative in the previous section. In particular:

- Fischer-Lichte theorised that performance comes into being by the bodily co-presence of actors and spectators, by their encounter and interaction; Gardair makes this plain in the practical work done in constructing and sustaining this performance situation.
- Fischer-Lichte speculates upon meaning-making; Gardair deploys pragmatics of talk- and gesture-in-interaction to demonstrate some of the specific processes by which a shared construction of context and meaning is achieved.
- Fischer-Lichte characterises performances by their eventness; Gardair describes how careful shaping of group dynamics achieved something mutually recognisable as event.

Both these studies gained insight into performance by attending to the interactional dynamics of the surrounding event. In both accounts, interaction is constitutive of the event itself and interactional dynamics between those present transform the experience for all. Interaction as the *liveness* of live events, then? Counter to this suggestion, much of the debate around liveness is focussed on no interaction. The next section will reconcile this.

1.2.3 No interaction?

A concept of liveness based on the role of interaction is at odds with historical accounts of the term and contemporary theory.

As first coined, *live* was established in opposition to recording. A performance was now classified either live or recorded, a distinction that was self-evident before radio and did not need to be named. Yet *live* events, as understood so far, existed before 1934, and would appear to have nothing to do with recording. Contemporary theory has the term introduced

to valorise the idea of co-present performers only when there was no material difference to the audience, and posits this valorisation persists to the present day (Auslander, 2008, p.59-63). Performance events feature co-present performers, it follows their promotion would benefit from a term valorising that: hence *live* events. There is no role for interaction, and only a tenuous connection to material fact or anything experientially distinct. But can this historically grounded valorisation explain how the *liveness* of theatre performance is often presented as central to its definition, and has increasingly become seen as something consumed; a commodity specifically purchased by audiences, even? (Reason, 2004)

In one sense, liveness is something that can *only exist within* a culture of reproduction. "The ancient Greek theatre, for example, was not live because there was no possibility of recording it." (Auslander, 2008, p.56). This point is made in the service of an exhaustive reappraisal of theories that privilege the live as real, authentic and timeless. However Auslander shows little interest in exploring what might be common to the experience of the theatre of ancient Greek and today. Instead, he argues that we now produce and experience performance as part of our experience of media. As Reason (2004) notes, even if the valuation of the live experience is a phenomenon created by the non-live, this does not negate the potential impact of audience perceptions of the live nor suggest what that experience of liveness might be.

Alternatively, liveness is something that has been *made visible* by the technologies of reproduction. This sense can be seen in Phelan's influential definition of performance as representation without reproduction: "Performance's life is in the present. Performance cannot be saved, recorded, documented, or otherwise participate in the circulation of representations of representations: once it does so, it becomes something other than performance. To the degree that performance attempts to enter the economy of reproduction, it betrays and lessens the promise of its own ontology." (Phelan, 1993, p.146). However this says little of what actually those qualities may be; though note in Phelan's vivid discussions of theatre's affective impact on its audience, she attends to the presence of the actors' bodies throughout. A liveness that is made visible by reproduction should not require reproduction for its definition. Better would be a theory of *being there live* worked up from first principles at live events, and so in principle be able to apply to other contexts – including where recordings were not available.

Phelan and Auslander's argumentation has resulted in assertions or denials of a fundamental opposition between live and mediated forms in order to prove the cultural superiority of one over the other (Fischer-Lichte, 2008, p.69). Both can be read as responses to a sense of technologies of reproduction having transformed or destabilised notions of performance, presence and the 'real'. These arguments can be found in Benjamin's (1936) writing on film, photography and prints having destabilised an sense of art that had held since time immemorial. In Phelan's valorising of a presence of the live body, there is a

parallel to Benjamin's arguments on authenticity of, and presence of, art. In Auslander's consideration of changing historical contexts, and "sensory norms" with that, he cites Benjamin's thinking as remarkably prescient (p.37). Auslander takes a very different position on presence, more aligned with continental theories of presence typified by Baudrillard's (1994) view that mass media have led to our discourse with the real becoming a discourse with the represented image, creating a cultural habit of reading the representation as if it were present.

A sense of convergence or confusion between the real and the virtual, and equivalently, the live and the recorded is accepted by Dixon (2007, p.127) when discussing liveness in his encyclopaedic review of technological performance. Causey (1999) neatly surmises the concern with the suggestion that theatre, performance, film and new media studies share similar concerns about the collapse of the real into the virtual and how identity is constructed in that (p.387) but also claims live and media forms are "the same, only different" (p.384). Causey's "the same" can also be seen in Dixon's acceptance of convergence / confusion between the real and virtual. Causey's "only different" can also be seen in Dixon's subsequent assertion that the real and the virtual clearly still differ in perceptual and phenomenological terms. Quoting Robert Lepage, Dixon turns to mysticism: "the audience in a theatre room is very different from the audience to a film, because they actually change everything on the stage by their energy" (p.131). Barker's (2012, p.43) cross-discipline review of liveness abandons the arguments as unworkable in favour of a wider comment on the literature: "it seems that, even in the most critically astute work in theatre and performance studies, there is a heavy investment in the need to defend 'liveness'. It is a primary cultural value, a *sine qua non*, of proper artistic experience". It is understandable that performance practitioners and theorists have been concerned with the (in-)difference of live vs. recorded performance. Yet this debate appears surprisingly unproductive for the question of what might be qualitatively distinct about being there live.

So let's start again. As first coined in 1934, *live* labelled reproduction for the listener being simultaneous with the production of its source. Live programming is still understood in this technical sense. It is used in diverse ways, broadcast professionals using the technique to get behind managed façades, experience risk, witness the authentic performance, watch outcomes emerge, to 'be there' for world events, and gain privileged insight into an event (Barker, 2012, p.46) – though these 'live' moments may well be replayed later. In this use of *live* there is no role for interaction, and there is no description of material fact given simultaneity is not guaranteed. Live transmission is however suggested to offer connection to 'shared social realities' as they are happening (Couldry, 2004). From 1948: "There you have the strongest appeal of television. It is life while it is happening. That is why it is wrong to associate it with the cinema. It is a different medium entirely, although much of the presentation technique may be similar to that of films" (John K. Newnham quoted in Barker,

2012, p.45; though note contemporary views in television studies often invert this positivity, seeing liveness as pretence and speculating about harm to the audience (p.47)). *Liveness* has been argued to be this connection to life while it is happening. The media involved have changed over time – television substituting for radio, and more recently mobile phones and the internet – and the required degree of simultaneity is quite fuzzy, nonetheless for Couldry it is a stable term. Deller (2011) uses the term in this way, extending the argument to micro-blogging practice seen with Twitter; Highfield, Harrington and Bruns (2013) demonstrate how Twitter became an unofficial extension of a live event. Liveness in this sense concerns universal human needs, and while the original technologies of broadcast denied any potential for interaction, the subsequent technologies are typically interactive and shift the focus from performers to within the audience – e.g. Twitter as a side-channel to broadcast of live events.

1.2.4 No co-presence?

If liveness has been made visible by the possibilities of technical reproduction, then 'event cinema' is particularly salient. Alongside traditional film programming, cinemas often now also show events. Music in concert, sport in a stadium, comedy in a club, drama in a theatre: these have all been captured and transmitted into cinemas 'live'. Intuitively, this would seem a very different experience. No co-presence between actor and spectator, sitting in a cinema. Yet liveness seems central to event cinema. It is there in the title of the spin-off company of Britain's National Theatre devoted to turning their productions into cinema: 'NT Live'. It is there in the title of the dominant scholarly work on event cinema, Martin Barker's 'Live to your local cinema' (2012), and the topic of two of the seven chapters within. It is there in theatre blogs, prompting calls for a re-examination of liveness, desiring "a liveness 2.0, if you will" (Trueman, 2009). It is there in audience responses: amongst a variety of open-ended questions asked to event cinema audiences, the one concerning liveness had responses with the longest word-counts overall Barker, 2012, p.62 – although, just as interestingly, audiences and industry may have normalised somewhat, with 'encores' of once live-streamed shows becoming more commonplace (Arts Council England, 2016).

How can event cinema's near-liveness inform theory? These are live broadcasts in a sense; a narrowcast presented to a massed audience surely has many similarities. Live transmission as a connection to shared social realities aligns with the shared experience, above, with other audience members in the theatre being viewed. However beyond this the concerns of the broadcast literature do not seem relevant (e.g. around the motives behind live vs. prerecorded programming). Appeals to the corporeal and/or interaction in performance studies fail as there is no co-presence between performer and spectator – though note the event cinema audience are extra spectators, as the source productions are filmed with their in-situ audience. Indeed, Barker (2012, p.58) sees event

cinema breaching all boundaries in the literature on liveness, given the different genres of performance, aesthetic practices 'owing most to television outside broadcasts', the 'assemblage of expectations' that go with cinemas as venue, new senses of locality (from each cinema auditorium to i.e. provincial cinemas showing productions from around the world), that the presentations emphasise differences to the source in specific ways, and that audiences know those differences, know that this is different. Such crossovers make the 'cacophony of contradictory claims and theories' surfaced in Barker's review look 'mildly absurd' and so with no theoretical model to apply to the near-liveness of event cinema, Barker turns to audience research to see what can be revealed about how actual event cinema audiences understand and feel about the issue of liveness in these events.

For the cinema chains, arts production companies, national bodies mandated to increase impact, there is also a clear need to better understand the nascent audiences of event cinema. While predominantly concerned with market and demographic data, the Arts Council England (2016) evidence review highlights a study of opera audiences that suggests the 'being there' of opera in a theatre becomes 'like being there' in a cinema; liveness more associated with having a shared experience with other audience members, either within the cinema or in the theatre being viewed. While lacking in detail, the 'shared experience' prompts consideration of Reason's discussion of social-spatial environments (section 1.1.1, *Drama in a theatre*), and of Couldry's sense of liveness as connection to shared social realities (section 1.2.3, *No interaction?*). Further, for some the source event had a buzz the cinema event didn't: out of 234 participants, 33 mentioned some form of excitement/thrill/buzz to their in-theatre experience while only 3 did so for their in-cinema experience; responses included "a cinema screening of opera can never be as thrilling as a good production in the Opera House" and "if you are in a theatre, there is a kind of hairs on the back of your neck thing with really great singers, which I don't think you quite get in any recorded form" Wise, 2104, p.15. A sense of occasion is also separated out from a sense of liveness, allowing a more interactional view of liveness as developed so far, separate from a cultural appreciation of the event.

People find valued experiences in theatre and cinema presentations, but through different routes, and not everyone will react the same way (Barker, 2012; Wise, 2104). For some, it is understood in a very traditional way. Nothing can replace being there in the house. For others, that understanding is unimportant, willingly sacrificed for something new with these events: an experience of privileged access. This access at its most basic is about affordability and clarity of view, and extends to interpretation, inclusivity and cultural understanding – often helped by extra information screened as part of the showing, such as behind-the-scenes interviews and guidance from the directors on what to watch and listen for Barker (2012). This sense of privileged access is not relevant to the question of what might be qualitatively distinct about being there live, it is a status argument similar to the

posit made in section 1.1.4, *Liveness?* that a football game can be seen and followed much better on television than at the stadium. Nonetheless, liveness is an issue for these audiences, and a sense remains even in the cinema presentation.

Barker's analysis identifies five aspects of liveness in the audience responses. Buzz, as above, is one. Aligning with the findings in the prior sections, interaction is mentioned explicitly, a sense of enabling the performance to happen (or wanting to) is suggested, likewise is adopting appropriate group behaviour. These five dimensions of liveness are realised differently for different audience members. Event cinema can be seen to succeed or fail for an audience member according to how each of these dimensions are interpreted; Barker's characterisations of 'Experts' and 'Immersives' stand opposed in their preferences. These five dimensions and orientations to them are collated in table 1.3. At this level of detail, there is one particular finding contrary to the narrative so far. 'Expert' audiences, in this analysis the most committed to the importance of co-presence, do not like being brought too close to the performers. They prefer to decline what they feel is an invitation to close an involvement in the performance because they risk losing their critical capacity to stand outside it.

Overall, event cinema draws attention to the audience. "There is no audience like a live audience which is part of the experience" (participant quoted in Barker 2012, p.68), yet a co-present audience remains in event cinema, with its own buzz.

1.2.5 Taking up the challenge

Fischer-Lichte (2008, 2010) challenges those studying art and culture to develop theories of aesthetic experience adequate to art-as-event[3]. To meet this challenge herself, she builds the concept analysed in section 1.2.1, *Speculative interaction*, through a series of descriptions of performance. Those studying embodied interaction have a role to play in meeting this challenge; section 1.2.2, *Observing interaction*, showed how Gardair's empirical enquiry and focus on interaction often made concrete what was speculative in Fischer-Lichte's account. What is not established is whether Fischer-Lichte and Gardair's central idea of performance as an interactional achievement can be generalised.

Understanding audiences seems central. That live events are comprised of people operating in social-spatial environments has been a recurring theme. These people have to be analysed as individuals; the particular pleasure and intensity of theatre in part due to tensions within the audience in Reason's (2004) analysis, the different dimensions of liveness being realised differently for different audience members in Barker's (2012) analysis. But are audiences somehow more than these individuals massed together? Gardair's claim is that they are, through a collective sense of membership individual behaviours alter and

[3]Chronological note: this work dates from 2004, in the full study *Ästhetik des Performativen*, and ACTAS keynote. The full study was translated into English in 2008, and the keynote was published as a book chapter in 2010

Immediacy	This is very much bound up with the fact that the event is unfolding as it is being watched. It is fostered by any sense of being able to interact with performers or communicate responses to the event.
Immersive	Being there as it is all happening, wherever that happens to be and whatever is offered.
Expert	Being able to interact with but also critically appreciate the performance.
Intimacy	This involves feeling close to the performers and the action, perhaps even in some sense enabling it to happen, by the way in which one responds. It also involves sensing how performers are achieving their performances.
Immersive	Getting as close as possible to the performers, in order to become intimate with them.
Expert	Getting an appropriate closeness that will aid appreciation and evaluation.
Buzz	This involves having one's reactions heightened by the awareness that other audience members are also engaged, excited, moved, and sharing the pleasure of the performance with them.
Immersive	Being with the right others makes it as good as being at the event.
Expert	'Others' have to be appropriate in that they share the right ways of responding.
Expanding oneself	This involves knowing that one may or should go away from the performance added to, enlarged, perhaps changed by the experience.
Immersive	Extra information boosts closeness and intimacy – giving access to performers' sense of their purposes.
Expert	Extra information can cloud perspective, superimposing performers' perspectives on necessary critical judgement.
Being (in) the audience	This has a distinct element of self-referentiality, that one knows how to be the kind of person who knows how to respond to performances of this kind. One knows that one understands and appreciates the conventions appropriate to this kind of event.
Immersive	Learning how to be at and part of this kind of event, and thus how to participate.
Expert	Knowing how to call on experience and expert criteria to evaluate the overall value of the performance.

Table 1.3: *Dimensions of liveness: descriptions with oppositional 'Immersive' and 'Expert' orientations in event cinema audiences. Compiled from Barker 2012, pp.66,67*

group responses are possible. In particular, the buzz of an event cinema audience is both challenge to the idea of performance as an interactional achievement and reinforcement of the importance of audiences in live events. It suggests that to investigate the liveness of live events in terms of interaction is to study interaction within the audience as well as with a performer. It suggests that the investigation is of mass interactions.

There is more at stake than meeting Fischer-Lichte's challenge. Sections 1.2.3, *No interaction?*, and 1.2.4, *No co-presence?*, show liveness is a contested term; 'Auslander vs. Phelan' in the 1990s is notorious within performance studies, these two influential positions are irreconcilable and in themselves resist development. There is a challenge to be met in defining liveness.

An approach to defining the term can be appropriated from Fischer-Lichte's (2008) series of descriptions of performance. Each description is built upon the last, each providing a richer account of performance. The first description considers co-presence.[4] The next, a materiality of performance considering co-presence alongside an examination of ideas and examples of corporeality, spatiality, tonality and temporality.[5] The next, meaning-making considering the materiality of performance[6], and so on. Each description better meets her challenge; her goal is an account of aesthetic experience. But traverse this series in reverse, and each description now pares away some richness, describing what enables that richness. Each description then better meets the challenge of understanding what is generalisable across experiences of live events. Liveness can thus be conceived as the foundational description of such a series of descriptions.

These two challenges are related. Taking interaction seriously in meeting Fischer-Lichte's challenge is to engage with the question of generalisability. An account of liveness that can rise above the contested history will come from engaging with the question of generalisability. Understanding the role of mass interactions in live events should allow a foundational account of liveness.

1.3 Towards a perspicuous account of liveness

Liveness is ill-defined. The diversity of live events makes generalisations difficult. Historical contingencies of the term *live* make generalisations difficult. Theories associated with the term may not address what is qualitatively distinct about being there live, the driving question of this work[7].

Liveness need not be ill-defined. By reviewing the literature with a focus on what is qualitatively distinct about being there live common themes have emerged. Intensified experience, interactional effects, and so on. In doing so, work towards a perspicuous account of liveness has been done; the term established in section 1.1, *Live music, comedy and drama*, and theorised in section 1.2, *Theorising liveness*.

By defining liveness in interactional terms, a simpler, more clearly expressed account seems possible. It would relate to the idea of performance as an interactional achievement. It would concern the sense of others within the event, and the perceptual environment that can help or hinder that. Generalised patterns of (mass) interaction should be identifiable. This could better ground theories of performance, and guide the design of interventions in live events.

[4] Chapter 3 – Shared bodies, shared spaces: the bodily co-presence of actors and spectators

[5] Chapter 4 – The performative generation of materiality

[6] Chapter 5 – The emergence of meaning

[7] Notable in its omission for this reason is Auslander's mediatization thesis put forward in *Liveness: performance in a mediatized culture* (2008). Here, mediatized performance is performance that has been captured into another medium. However, this is not a study of mediation, an aspect this chapter has shown to be relevant, rather the capture is a cultural-economic one.

Chapter 2

Examining an event

In chapter one, the accounts of live events were from written sources. In this chapter, the account is first-hand: a stand-up comedy event is described and analysed. The idea of performance as an interactional achievement is supported, most simply by the performer stating the act is going to be a dialogue with the audience. Certain analyses from chapter one are then successfully elaborated using the new data, further supporting the idea of interactional achievement. The study of mass-interaction is limited however, and this chapter ultimately shows the need for a better understanding of what to look for in human interaction, and a different approach to data collection.

2.1 Observing stand-up comedy

To examine an event, an actual event has to be chosen. The choice should not matter, as the investigation is of the liveness of live events rather than the particularities of any one event or genre. But a choice has to be made, and of the examples developed in section 1.1, *Live music, comedy and drama*, the choice is stand-up comedy. This is a pragmatic decision, the rationale being the performance that needs to be observed and analysed is minimised – a single comic in a spotlight – while the audience responses are maximised – laughter is plainly apparent.

In section 1.2.2, *Observing interaction*, an ethnography of street performance events was examined and found an ideal event to look for evidence of the role of interaction in a live experience. For instance, in street performance the audience needs to be opportunistically constructed from passers-by, and this provides many opportunities for shaping the audience. Passers-by are more than potential recruits for the audience: interactions with passers-by are used to demonstrate to all that the space is a performance place; interactions addressed to passers-by but designed for the audience are used to construct certain forms of participation as welcomed while discouraging others. In street performance it seems difficult to elicit group responses from the crowd. The performers have to work at it, and in their failures and corrective action causal analysis starts to become possible. Performance as an interactional achievement was demonstrated, but it is an open question whether these findings apply in more conventional performance situations. For instance, what if there is already a mass of people spatially arranged to watch the performance, their commitment to watching demonstrated by, amongst other things, queuing in advance?

This section will look for evidence of the role of interaction in a stand-up comedy event, considering in particular Gardair's (2013) claim that through structured sequences of interaction (street) performers get people's attention, achieve the recognition that what is going on is a performance, build a collective sense of audience membership, establish moral obligations to each other and the performer, and train the audience how to respond. These ideas broadly structure the following examination of an event.

2.1.1 Ten minutes of Barry Ferns

Angel Comedy is a successful, nightly comedy club in London.[1] The observation was made on a Thursday night, advertised as *Angel Comedy RAW*: "Every Thursday, our wonderful resident MC BARRY FERNS presents some of his favourite comedians from the UK comedy circuit, as well as some very special guests and leads a night of the unexpected!

[1]The UK comedy website Chortle's 'Best small club' in their annual awards 2015 (Chortle, 2015); Leading UK listings website TimeOut states it is "consistently excellent", "gets absolutely rammed" and aggregates 5 out of 5 stars from user reviews (Time Out London, 2015)

It can be big acts – and some newbies too."

The club night has no fixed premises and minimal staging. It is held in a function room in a pub, rearranged with the tables pushed to the back so that the seats can be made into rows facing the stage. The stage itself is a raised platform in the corner of the room, ~2x3m, with a black backdrop hung to cover (and in contrast to) the period decor. The stage area has the acts well lit when the room's lights are dimmed, but lacks any particular theatrical effect such as the circle of a spotlight. The venue was full from the time of doors opening, with around 60 people attending. All seats were filled, people stood at the sides and back, and sat on the floor at the front to the side of the stage. No admission was charged. A sketch diagram is provided in Figure 2.1.

The performance was captured using a DSLR camera with external microphone, mounted in a fixed position on a tripod. The position was as per a standing audience member, to the side of the seated audience, around a third back. The captured aspect ratio is 16:9 with a horizontal field of view of 38°.[2] This captured the stage in fine detail but did not extend to the audience beyond a few heads of the first row. The view can be seen in Figure 2.2.

2.1.2 Displaying a performance

For around ten minutes from the doors opening, music plays from what appears to be a guitar amp and iPod, with the room's lights up. Then, offstage, someone picks up a microphone (on a stand, also offstage), turns down the music and starts the show, as per transcript excerpt 2.1 (see appendix A.1, *11–83s: "Hello, hello, hello."*). Thus announced, Barry Ferns then performs for 10m15s, after which he introduces the next act and leaves the stage.

2.1

```
Barry(anon): Ladies and Gentlemen are you ready for
             a night of stand-up comedy?
Audience:    <claps><cheers>
Barry(anon): If you are, gimmie a yeah
Audience:    <yeah>
Barry(anon): Gimmie a whooh-yeah
Audience:    <ooh-yeah>
Barry(anon): Very good. Please welcome on stage –
             a round of applause
             – your MC for the night Mr Barry Ferns.
             <steps on to stage>
Barry:       It's just me there is no-one else it's just me.
             <house lights dim>
             HELLO!
```

[2] A 35mm focal length lens was used with a APS-C sized sensor, a combination accepted in photography as reproducing a field of view that appears natural to the human observer, though note this is less than what someone standing there would see. See discussions of 'Normal Lens', such as https://en.wikipedia.org/wiki/Normal_lens

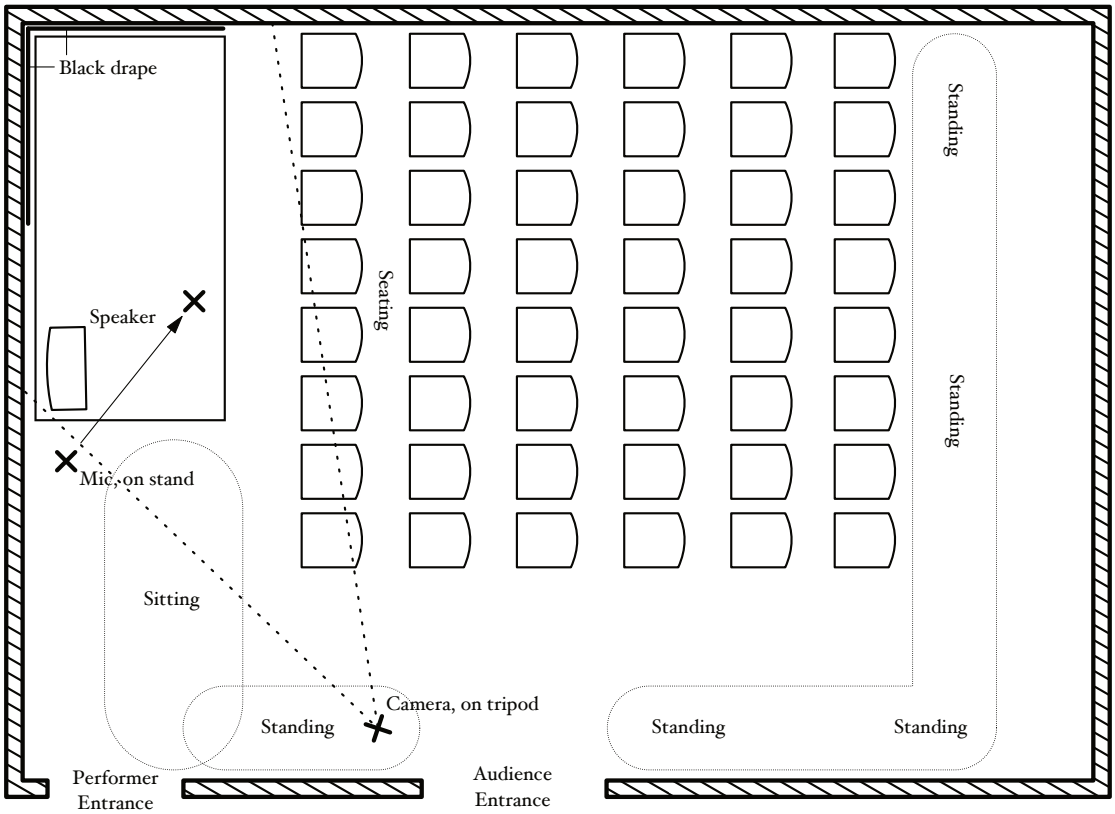


Figure 2.1: *Sketch staging diagram of Angel Comedy.*

Figure 2.2: *Frame from Angel Comedy capture at 11s with annotations. Barry Ferns is turning down the music and introducing himself.*

The most obvious point is that interaction is explicitly encouraged in this, the opening announcement. The interaction is clearly a critical element in shaping the live experience here. But why does Barry choose to announce himself in the third person? More straightforward would be to enter the stage and start the proceedings directly, e.g. "Welcome, I'm Barry Ferns, your host for the night". One answer could be that larger comedy clubs often have an announcer role separate to the compère, and Barry announcing himself is an appropriately 'cheap' approximation of this (the night is free, after all). Perhaps so, and a joke is made from it. But this does not answer why have an announcement at all? The sequence warrants further examination.

The sequence opens with a question, explicitly posed to the crowd. A gesture at the end of the utterance reinforces this, with Barry's gaze scanning across the audience. The crowd responds with claps and cheers, and recognising this Barry smiles and continues with a request for a specific response: "if you are, gimmie a *yeah*". Barry then looks, straight-faced, at the audience. A 'yeah' response comes quickly, and with conversation-like timing Barry continues with "gimmie a *whooh-yeah*". This utterance is more affected, leaving Barry's face slightly contorted. The audience response comes again with conversation-like timing and Barry immediately starts smiling. His head then nods as his gaze appears to look at two adjacent (off camera) audience locations, stage-right. The impression is of satisfaction, reinforced by a gesture of flashing his eyebrows. This apparent satisfaction is then turned to reward for the audience, with the utterance "very good" being made to the latter presumed audience member, before returning his gaze back to – perhaps slightly above – the central audience.

This behaviour is accountable in the ways street performers make their actions recognisable as the pre-cursors to a performance, as discussed in section 1.2.2, *Observing interaction*. Gardair (2013, pp.87-110) shows that in street performance, to engage passers-by what matters is how actions are designed for them and how their responses are integrated into the act. Barry starts with a room full of people waiting for the show to begin, so is clearly in a different position to street performers who have to opportunistically assemble a crowd from passers-by. They are, however, conversing amongst themselves. Barry clearly needs to get people's attention, and as Gardair shows achieve the recognition that what is going on is a performance.

Barry succeeds in getting people's attention, demonstrated by the chatter dying down over the 12s of this sequence. The above description has many features that emphasise how Barry's actions are designed for the waiting crowd. Achieving the recognition that what is going on is a performance seems explicit in what happens next. The above takes place off-stage, to the side. The microphone stand is also here. Following straight on from the "very good", Barry's gaze disengages from the audience slightly and excitedly shouts "Please welcome on stage – a round of applause – your MC for the night Mr Barry Ferns."

Barry's eye-gaze then reconnects with the audience while making a gesture that raises his head up with microphone tracking his open mouth. This seems to have two purposes. One is to build anticipation, as if he is holding his breath. The other seems to be signalling to the audience their expected response, screaming harder (c.f. section 2.1.4, *Training the audience*). At the climax of this head-rise, Barry drops his head, and says "It's just me there is no-one else it's just me" while stepping onto the stage. This is done sideways to the audience, and he brings the mic stand with him. Barry turns to face the audience, places the microphone stand, and without pause says a loud, sharp, "Hello!". This punctuates the proceedings, and these actions make clear this is now Barry Ferns, a performer, performing. Moreover, this minimises any crossover with the actions of establishing that performance, and so association with the non-performer person. Before this point, he was anonymous and his body was spatially de-emphasised. The transition is further emphasised by the house lights dimming as Barry comes on stage.

Gardair's analysis draws attention to the performative use of objects, which seems relevant considering that having the microphone stand already on stage waiting would, amongst other things, have simplified getting onto stage. The stand is not integrated with an addressed gesture, as in many of Gardair's examples. Nonetheless, it is handled throughout Barry's warm-up act, and then remains on stage for the subsequent acts. It is introduced at the moment the performance is established and the association with performance then further established during the act. In doing so it gains meaning for those present it would not have had as preplaced object.

The idea of performance as an interactional achievement is supported by this opening sequence of the night. Barry's choice of indirection as announcer and use of objects fits the analysis Gardair (2013, pp.87-110) demonstrates in how street performers need to display a performance, as the first stage of the interactional construction of the audience.

2.1.3 Shaping the audience

Following Barry's now-introduced "HELLO!", he continues with many more, variously addressed to "everybody", "people at the back", "people at the side" and "people who have seats". Cheers come in reply from the crowd – notably more cheers than "hello". They are loud and timed together (the totality of participation is hard to gauge given the limitations of the recording, but while the odd individual's response can be picked out, the overall effect is of a massed cheer). By the end of this sequence, there is no more chatter amongst the crowd. The beginning is excerpted in 2.2; full sequence in appendix A.1.

2.2

Barry:	HELLO! Hi. Hello everyone. HELLO!
	<puts mic in stand, holds stand with left arm>
	Marvellous marvellous people. Hello, hello, hello.
	And hello people at the back.
	<right arm gesture with open palm to back>
	How you doing at the back?
Audience:	<cheers>

The most obvious point is that interaction continues to be explicitly encouraged. Again, the interaction is clearly a critical element in shaping the live experience here. But why not, as the subsequent acts did, open with a single hello or similar brief acknowledgement of the performance situation?

The extended "hello" of this sequence is also accountable in ways touched upon in section 1.2.2, *Observing interaction*. "Hello" in a social situation brings expectations of behaviour. Gardair's (2013) argues that street performers exploit this projection of social engagement, knowing that people are sensitive to not providing a response. She shows it is a way for the performer to establish their expectation of a responsive audience before a social situation has changed into the performance situation (p.114). Barry's behaviour is accountable in this way. The newly displayed performance intrudes onto a social situation, seen explicitly in the many conversations amongst the crowd that have not (yet) stopped. This makes it appropriate for Barry to impose those expectations onto the crowd. The non-social "hellos" that come later suggest he is indeed successful in creating a performance situation from the social; for subsequent acts an opening "hello" appears a simple acknowledgement of the performance situation, an announcement that they're starting with no "hello" back from the audience.

Why does Barry addresses his "hellos" to different parts of the room? In Gardair (2013) analysis, for people to follow the performer's direction, each member of the audience must feel as if they are being personally addressed. Generic addresses are only effective when backed up by gestures which identify specific parts of the audience (p.127). This suggests 'And hello people at the back', accompanied by a right arm gesture, and what follows are pragmatic actions Barry takes to engender a response from the audience. But also Gardair's wider argument emphasises group behaviours and how membership of groups is built. This extended 'hello' can be interpreted in this way. Competition is established between parts of the crowd to display their audience-ness and contrasts are made between parts of the crowd; both techniques fostering group identity. Starting Barry's performance, the crowd is, prima facie, not yet an audience. At least, it has not yet a chance to exhibit audience-like behaviour such as cheers or boos in response to the performance (c.f. section 3.3, *Audiences?*). In the exchange of "hello" greetings, the crowd gets this chance.

2.3

```
Barry:      Allright. Ok. Hi! Welcome to Angel Comedy.
```

The formal welcome to the night and Barry's facilitation duties as compère are deferred for 20s with the announcement and then by a further 50s of crowd-work in the extended hello detailed above. Barry only states the welcome once satisfied his performance is established, and an audience is established from the crowd.

The idea of performance as an interactional achievement is supported by this sequence. Having achieved the recognition that what is going on is a performance, the next goal in Gardair's analysis is to build a collective sense of audience membership and establish moral obligations to each other and the performer. The specific actions taken here, and what Barry defers to do so, fits this analysis.

2.1.4 Training the audience

Barry establishes that interaction is an integral part of the night throughout this performance. As part of the "welcome" speech that follows the extended "hello", Barry describes the night to come while simultaneously soliciting the audience for a response. This is shown in excerpt 2.4, which comes 40s after the "welcome" of excerpt 2.3; full sequence in appendix A.2, *115–141s: "Wooo"*. With the line "Good response..." Barry makes the interactional context explicit.

2.4

```
Barry:    on our Thursday night ladies and gentlemen,
          is our new act, <arms sweep back> new material night.
          <arms slowly rising up, palms out to audience…>
          ooooooooo
Audience: <pause> wooooooo
Barry:    <arm gesture end><smiles><nods>
          Good response to me going <starts to repeat gesture>
```

Another 40s or so later, Barry is focussed on his role introducing each act through the night. He is again explicit about interaction shaping the experience of the night. This is shown in excerpt 2.5; full sequence in appendix A.3, *167–247s: "Go Red Sox!"*. He foregrounds the audience as part of the night. He says he will be chatting with them. He calls out the role of eye-contact in facilitating the chat.

2.5

```
Barry:      And we've got you guys. And um, I'm curious
            as to... I'll be chatting with you guys as we go on.
            The people in the front row are now going "oh shit,
            don't make eye contact. Don't make eye contact".
```

By the point Barry should be introducing the first act, Barry solicits audience responses and readily receives coordinated, massed responses. This is shown in excerpt 2.6; full sequence in appendix A.5, *594–625s: "Welcome on stage to..."*. With the line "I like that...", Barry again makes the interactional context explicit.

2.6

```
Barry:    Are you ready for your first act of the night?
Audience: Yeah
Barry:    If you're ready for your first act of the night,
          gimmie a 'yeah'.
Audience: YEAH
Barry:    Gimmie a 'fuck off Barry'
Audience: Fuck off Barry
Barry:    I like that, that's as engaged as you've been.
          But I will, I will.
```

Barry finishes by introducing the first act of the night. This is shown in excerpt 2.7, completing appendix A.5, *594–625s: "Welcome on stage to..."*. The crowd applaud, vociferously. While Barry has been explicit that interaction is integral to the night, is such rhetoric sufficient to explain the audience response?

2.7

```
Barry:    Please start the applause
          <both arms rise up with palms up and fingers wiggling>
Audience: <applause...>
Barry:    bring it up bring it up bring it up
Audience: <...applause...>
Barry:    bringing it back to you
          and welcome on stage to SAM RUSSELL.
Audience: <...applause, cheers>
```

(Gardair, 2013, pp.128-148) argues that having built a collective sense of audience membership, performers then train the audience how to respond. For the street performances studied, she shows that the performers cannot rely upon rhetorical devices and large gestures alone to systematically elicit responses that are a) the expected response and b) produced in the expected tempo and with the expected magnitude. To compensate, they explicitly teach their audience when and how to contribute. Notably, Barry's final gesture closely resembles a gesture Gardair identifies street performers as using. Figure 2.3 shows Barry performing the gesture, which takes place during excerpt 2.7.

Gardair shows how features of this gesture can be taken advantage of in an interactional context. She shows street performers establishing a gesture-response pair through a particular sequence of steps. First, the gesture is used in a greetings adjacency pair, "hello-hello", taking advantage of the projection of social engagement discussed in the previous section. The gesture is clearly demonstrated as part of the performer's opening turn, and

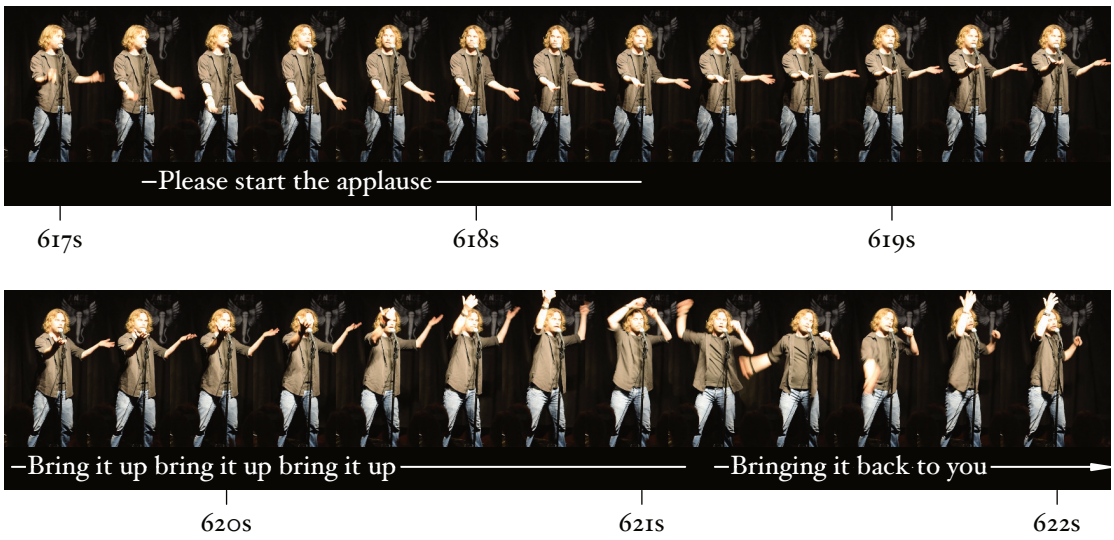


Figure 2.3: *Barry Ferns soliciting applause just prior to introducing the next act.*

then repeated for the projected turn of the audience. The gesture starts by drawing the hands back, which then turns sharply into sweeping the hands upwards. The performers' speech pauses during this initial drawing hands back phase, and the "hello" utterance is simultaneous with the sweeping upwards motion. The audience response is then judged in a follow-up comment, transforming the pair into a three turn sequence comparable to classroom appraisal (Tsui, 1989; Lee, 2007). The critical aspect is that in the third turn the performer (i.e. teacher) acts on the audience (i.e. student) response in a way that should affect future interaction (e.g. "very good"); contrast that appraisal with a reply of gratitude, "very good" would encourage that specific response to be used again in a way that a simple "thank you" wouldn't. This 'hello' three turn appraisal sequence is repeated, until the gesture need not be demonstrated in the performer's 'hello'. At this point, the initial drawing back motion of the gesture to accompany the audience 'hello' starts before the end of the performer's 'hello', such that the upwards motion starts at the moment the audience would respond with conversational timing. The initial drawing back motion is a preparatory signal for the audience, and in the back then up motion a mechanism to control the timing of the audience response. This timing function can further be seen in the end pose with hands fully raised being held for the duration of the expected response, and dropped sharply when the response should stop. The repeated three turn appraisal sequences makes this clear both for audience and analyst. Once the gesture-response pair is appraised satisfactorily, the performer introduces the expected response. Further three turn appraisals happen until the gesture alone is sufficient, with no request in the speech. Gardair speculates that teaching this non-spontaneous response to the audience, distinct from applause or cheers, makes the audience aware of their own actions and accentuates that they, as a group, are distinct from others.

Barry's performance features these techniques. The final gesture is established earlier,

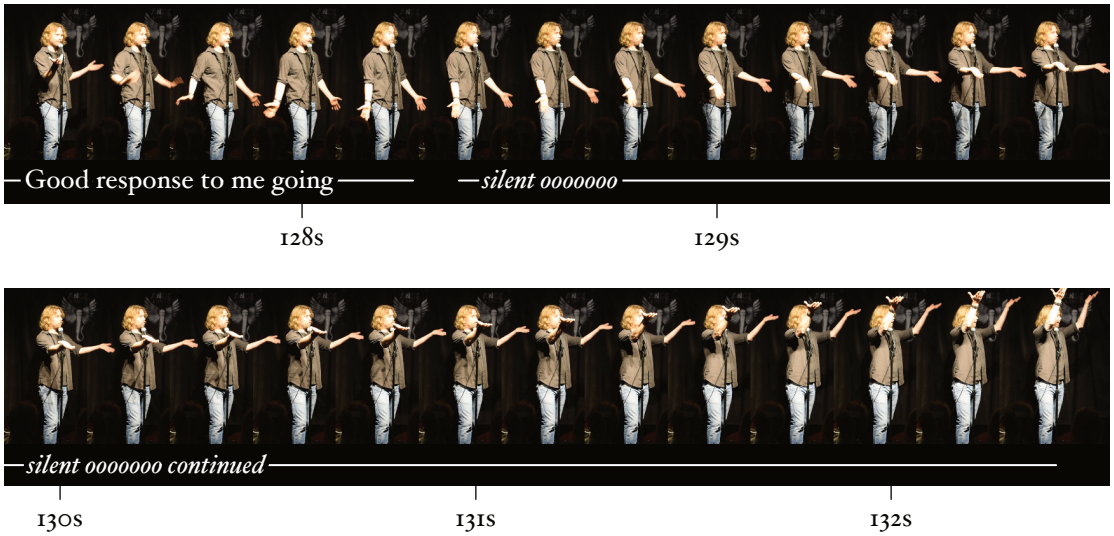


Figure 2.4: *Barry Ferns soliciting a massed "oooooo" from the audience.*

accompanied by a three point appraisal. This earlier use is shown in Figure 2.4, part of excerpt 2.4. "Hellos" are used to earlier establish a punching gesture, which through his performance signals every projected audience turn, and in the sharp motion followed by hold and release, attempts control over its timing. This starts in excerpt 2.2 and through the full sequence in A.1. The end gesture is in fact a combination of these two gestures, the punch projecting renewed cheers – see it accompanying "bringing it back to you" in Figure 2.3.

This analysis of ten minutes of Barry Ferns firmly supports performance being an interactional achievement. Most simply by the performer stating the act is going to be a dialogue with the audience, but more pervasively in the elaboration of Gardair's analysis of the needs and techniques of street performers to a new performance context. Barry's act might be described as "warming up" the audience, and perhaps this is an intuitive articulation of the idea of performance as an interactional achievement.

2.2 Motivating an interactional focus

The critical element in shaping the live experience in this case is the interaction. Looking at what actually happens in the event makes this obvious. Interactive processes involving rhetoric, gesture, gaze and selective address have been identified, and the performer has provided clear and explicit evidence by directly commenting on them. This is the start of how defining liveness in interactional terms should lead to a simpler, more clearly expressed account. Contrast with the problems in defining the value of liveness in socio-cultural terms (section 1.2.3, *No interaction?*), or valuing liveness merely as audience participation, and then rejecting even that as pretext (section 1.1.4, *Liveness?*).

Simpler is not *simple*, however. Following Fischer-Lichte's concept of performance and

Reason's idea of social-spatial environments, the argument for a generalised, interactional account of liveness lies in pervasive, subtle interaction. However the data here is inadequate for study of the interactional cues of everybody present, throughout the event. The analyst needs to look into the dark of the audience, and needs to look for cues of interaction that go beyond direct address. The study of this will require a better understanding of what to look for in human interaction, and an approach to data collection suited to the auditorium and the phenomena identified. The review coming next in chapter three addresses the gap in knowledge, and the experiment presented in chapter four develops such an approach to data collection.

Chapter 3

Audiences and interaction

The observational study of stand-up comedy in chapter two demonstrated a gap in theory, method and instrumentation for the study of mass-interaction in live events. Chapter three addresses this gap by reviewing literature on co-present human interaction.

Applying the concerns of this literature in the context of live events leads to consideration of mass spectatorship. Is there a distinction between people who are merely massed together and 'audiences'? This chapter argues that there is and that it consists in the specific kinds of social organisation involved.

3.1 Dynamics of interaction

3.1.1 Dialogue

Human interaction is fine-grained and pervades everyday encounters. One demonstration is the interactive construction of natural conversation. The sentence "I gave up smoking cigarettes one week ago today actually" is coherent and appears pre-formed. Yet Goodwin (1979), who recorded its telling at an informal dinner, demonstrates that this sentence was constructed in a highly interactive manner. In the course of its production the unfolding meaning of John's sentence is reconstructed twice, a new segment is added to it, and another is deleted prior to its production but replaced with a different segment. This is shown by applying a set of rules developed through analysis of a wider corpus to the actual utterances and non-verbal behaviour recorded in this instance.

3.1

```
John: .................Don,,.........Don_______
                        [             [
      I gave, I gave u p smo king ci garettes::.=
                                [
Don:                   ........X_______________
Don: =Yeah,
                       (0.4)
John: ......Beth___________,,.......Ann___________
            [                       [
      l-uh: one-one week ag o t'da: y. actua lly,
                              [                 [
Ann:                   ........Beth____,.........John
```

Compare the neatly extracted sentence with the transcript sufficient for Goodwin's exposition (3.1, excerpt compiled from Goodwin 1979, 1981). The detail of the transcription system need not be understood to get a sense of the fine-grained structure of what might otherwise be taken at face-value.

The analysis considers how the utterances and gaze of those present are sequentially organised and timed. First take "I gave". In the sentence form this is two words of the subject-predicate-object construction 'I gave-up smoking'. Yet as spoken, there is a repetition despite this making no contribution to the meaning of the sentence. It did make a contribution to mutual engagement, however: the addressee was shown by the speaker's gaze, and the addressee's gaze back was secured by the repetition. This can be seen in the transcription of gaze, where the gaze is marked as the first line, with any utterances below. Looking at the gaze transcription further it can be seen that over the construction of the sentence, its speaker, John, gazes at three different individuals. When John begins the sentence, he looks for an addressee for whom the announcement would constitute new information: his friend Don rather than wife Beth. The gaze changes to Beth on the halted

"l-", which probably was the beginning of "last week". Instead, the emerging sentence is now re-tailored for Beth: highlighting that the occasion marks an anniversary of him quitting – new information for her. What started as news is now reframed as the announcement of an anniversary, all the while remaining suitable for all at the table. To conceptualise the creation of this sentence as the speaker first thinking it up and then speaking it out would clearly be a mistake. To analyse this sentence in isolation from the interactive process that created it would miss so much.

In conversation, speakers are highly responsive to those around them. Speech and gesture are adapted moment-by-moment to manage attention and maximise relevancy. Through this example and others, Goodwin (1979) shows the adaptation happens to the degree that multiple reformulations can happen within one speaker's turn, noting the fact that coherent sentences emerge, and that they were apparently the sentences being constructed all along, to be one of the more striking features of the process.

Joint projects

The essentially interactive nature of dialogue is highlighted by the *collaborative model* (H. H. Clark, 1996). The rationale is this. If conversation is a social process, then people are speaking for the benefit of the other participants and they should make them themselves understood to everyone. But how do they know whether they have been understood? They can only assure themselves that others have understood what they meant. To assure themselves, all they have to interpret are the actions of the others. For the actions of others to become signals of understanding, the parties must coordinate to establish the mutual belief that they have understood one-another.

Describing this process, H. H. Clark and Schaefer (1987) contrast dialogue with literature. In literature, the content is presented, and that is that. In dialogue, however, there is a further phase of acceptance: only once what was presented has been accepted, has the contribution to the conversation been made. This is characterised as a process of grounding, so called as the participants try to establish common ground between them. In order to progress the common ground, the interactive process of grounding must pervade dialogue.

How does this process manifest itself? To collaborate, people have to work within particular constraints. Conversation happens in the moment, sustained by the effort of those conversing. If an utterance is missed by a listener, the opportunity to hear it is gone. Continuous evidence of attention and understanding is essential to avoid losing context as the conversation proceeds, as the task of reconstruction would become increasingly burdensome. Thus with each of their utterances, speaker and addressee should collaborate to ensure that the addressee has understood the speaker 'sufficiently for their current purposes' (H. H. Clark & Wilkes-Gibbs, 1986). Consider the following transcript of a brief call to directory enquiries, one of 602 calls analysed where a number was given (3.2, H. H.

Clark and Schaefer 1987).

3.2

```
Operator: Directory Enquiries, for which town, please?
Caller:   In Cambridge
Operator: What's the name of the people?
Caller:   It's the Shanghai Restaurant, it's not in
          my directory, but I know it exists
Operator: It's Cambridge 12345
Caller:   12345
Operator: That's right
Caller:   Thank you very much
Operator: Thank you, good bye
```

The operator presents the number in the conversational turn `It's Cambridge 12345`. 91% of callers then demonstrated mutual acceptance in the following turns, as per the collaborative model. In 68% of these calls the next turn presents the number back, here `12345`. This displays full hearing. Per the collaborative model, the operator should then indicate the correctness of that display, ensuring mutual acceptance. 89% of the time, this was done explicitly with the next turn being one of `right`, `that's right`, `yes` or `ok`. 11% of the time, the acceptance is implicit by initiating the next contribution without raising any error. The strength of the evidence comes from not just the prevalence of this pattern of mutual acceptance, but of the model's ability to account for the less common structures too.

Displaying full hearing was not the only response to their situation. By saying `Thank you` they would have presupposed full hearing, the next most common response (11%). By saying `12?` they would have presupposed incomplete hearing (9%). Or `I didn't hear you`, asserting no hearing; `What did you say?`, presupposing no hearing; `12345?`, presupposing fallible hearing; `Right`, asserting full hearing. Each of these responses should start the operator and caller down a different path. Applying rules such as *strongest initiator* to these paths taken, the predications of the collaborative model correspond to this data (see H. H. Clark's (1996) 'strength of evidence' principle).

The evidence for Goodwin (1979), above, and H. H. Clark and Schaefer (1987), here, comes from study of a corpus; 50 hours of video recording of social gatherings in the US and 17 hours of recording of telephone calls to directory enquiries in the UK respectively. Rules are inferred that predict the observed phenomena consistently across that corpus. The theoretical models and inferred rules may however be bound to the specifics of these corpora; in particular, it is perhaps not surprising to find much checking of information in calls to directory enquiries. Something more contrastive that could test these models and rules is needed.

Interacting, or not

There is also experimental evidence that this process has consequences for addressees' understanding. Schober and Clark (1989) contrasted listeners who could interact in the conversation and those who could not. Taking the view that conversation proceeds contribution by contribution as a process of grounding, they reasoned that manipulating a listener's ability to participate in grounding should affect their understanding of what the speaker is saying. An addressee would be an interacting party, able to coordinate their understanding. An overhearer would not be an interacting party, unable to give feedback and so coordinate their understanding. The overhearer would thus be at a disadvantage. They set about to test this contrast.

An experiment was designed based upon a communication task known in psychology literature. Two participants were to match a series of shapes. Without being able to see each other or their materials, the *director* is tasked with instructing the *matcher* to arrive at a given arrangement as quickly and accurately as possible. Schober and Clark's variation was to record each conversation and then repeat each series of trials with a matcher listening to (and even controlling playback of) this pre-recorded conversation. In this variation, each trial of shape arrangement is performed in a condition where the listener is an addressee and in a condition where the listener is an overhearer.

As expected from previous stagings of the experimental task, each director-matcher participant pair followed a pattern of collaboration over a series of given arrangements. The directors first described each shape, and subsequently referred to them with definite descriptions, which got shorter and shorter as each pair progressed through their series of arrangements. These references included perspectives offered by both the director and the matcher. The negotiation around these references also reduced as each pair progressed through their series.

The novel data came from repeating the task with the listener role as an overhearer, and indeed there was a contrast. Directions that before had led the addressees to achieve near-perfect accuracy now led overhearers to mediocre results. The primary measure is accuracy: did the matcher choose the shape the director was describing? The addressee-matchers averaged 93% accuracy on the first trial and were perfect from the fourth trial onwards. The overhearer-matchers averaged 81% accuracy on the first trial and by the sixth trial had only increased to 95%. Overall, addressees achieved 99% accuracy compared to overhearers' 88%. Moreover, overhearers who joined a series late, at the third trial, performed even worse, increasing only to 70% by the sixth trial. Directions that had led the addressees to place their shapes – more or less the point of having chosen – when the direction stopped now led to the overhearers to place their cards with much less synchrony. 20% were placed earlier, and an early placement saw a slightly better accuracy. 31% were placed later, and a late placement saw much worse accuracy. Qualitative analysis of the overhearer's speech

shows them verbalising their difficulties. If they had been addressees, they would have had the opportunity to resolve these in the ongoing conversation.

The experimental results show that witnessing the build-up of common ground has very different outcomes to participating in it: to be an interacting party is to have a demonstrable advantage over someone who is not. The disparity couldn't have been due to the quality of the speaker's directions, as these were constant between conditions. Nor could it have been listening more carefully, as this again was constant between conditions, and any individual effects averaged out. Neither the ability to pace and review a recording, nor some quality of listening to mediated/unmediated speech, were found to change the disparity: the results held when instead of recordings being used the speaker, addressee and overhearer were co-present and the experiment run 'live' (though, as per design, still unable to see each other).

When analysing within or across utterances in conversation, interactive processes can be seen to shape so much of what is there. While the coherence of the produced speech can mask these dynamics of interaction from everyday experience, this interaction is well explained in the literature. Cues of interaction can, and have been, identified through fine-grained observation of verbal and non-verbal acts, and principles to model and motivate that interaction can, and have been, derived for which subsequent testing bears out.

3.1.2 Beyond the dyad

Conversations have structured audiences. Addressees and overhearers have been contrasted. Within a dyad, the structure is straightforward: the audience member is one or the other. But for larger audience sizes? Does an utterance carve the world beyond the speaker into precisely these two parts? Goffman (1981) argues it rather opens up an array of structurally differentiated possibilities[1].

Goffman frames conversation as "ventures in joint orientation", where "two or more persons in a social situation can jointly ratify one another as authorised co-sustainers of a single, albeit moving, focus of visual and cognitive attention" Goffman, 1964, p.135. This creates a boundary that can be explored, with results that can partition the audience in terms of the mutual access they have to the speaker (Goffman, 1981). *Byplay* has a subset of ratified participants conducting a further conversation, subordinated to the main (i.e., timed, and pitched to constitute a perceivably limited interference). *Crossplay* is similar, but between ratified and non-ratified participants. *Sideplay*, between non-ratified participants. Further, non-ratified participants may have deliberately engineered their outsider situation, and so be *eavesdroppers*. Or they may merely be *bystanders*, likely doing what they can to "maximally

[1]The broader point was already established: "Long traditional in our culture is the threefold division between speaker, hearer, and something spoken about. It has been elaborated in information theory, linguistics, semiotics, literary criticism, and sociology in various ways. [...] All such schemes, e.g. appear to agree either in taking the standpoint of an individual speaker or in postulating a dyad, speaker-hearer (or source–destination, sender–receiver, addressor–addressee). Even if such a scheme is intended to be a model, for descriptive work it cannot be." — Hymes, 1974, pp.53–54.

encourage the fiction that they aren't present" (p.132). And when an attempt is made to conceal any of this subordinate conversation, it becomes *collusive*. This instantaneous view known as *participation status* – the relation of any one person to an utterance – fits phenomena such as people surreptitiously editorialising on what they are overhearing into a typology (*collusive sideplay* in this case). The typology can describe complex audience structures with speaker and one subset of the audience dealing with the talk in one way, while another subset attends to what is being said but uses it as a point of departure for its own talk.

How are these audience structures created? Do they need to be maintained? In proposing the participation status typology, Goffman (1981) remarks that the ratified participant boundary will be be unstable, with conversations starting and stopping, participant groups splitting and joining. Similarly, he remarks that speakers may police their listenership to guard against eavesdroppers, or to bring back strays and encourage incipient joiners. This suggests that any such audience structure is the product of a dynamic process, however these observations are not developed. Detailed evidence comes instead from a backyard picnic in mid-western USA, where it is perhaps expected that people will join in, and as social peers jockey for their view to be accepted.

An audience and the interpretive work in which it is engaged are shown to be constituted through a dynamic process of ongoing interaction by Goodwin (1986). The approach is similar to the earlier work at the dinner table ("I gave up smoking...", Goodwin 1979), the analysis demonstrated this time using a few minutes talk amongst five participants. In brief, a report of a sporting incident is told by someone for whom the events are serious and dramatic. One listener has a similar alignment, another has a stake but is a non-expert, and of these three male participants discussing a male dominated sport, two have female partners present. The main teller's interpretation of events is challenged by the non-expert, and combined with exclusion of, and indifference from, the partners, the main teller has narrative trouble at the climax of the story. He faces the task of countering the non-expert and getting others to listen to and possibly accept that counter. And beyond that, finding recipients who will continue to act as an audience to his talk. The following transcript extract covers part of this trouble (3.3, Jefferson transcription system, Goodwin 1986).

Mike is the storyteller. As this transcript starts, Mike is building to his climax. Gary is the non-expert. Gary interjects with a proposal that the incidents being described are "All show" and the storytelling as something that has now reached its projected climax. These two options are actualised by different participants; Carney's actions align with Gary's proposal, belittling the characters in Mike's story, while Curt continues to act as an audience to Mike's talk. This is seen non-verbally too, where having made his 'story over' proposal, Gary withdraws from Mike's talk and asks for a beer. Curt, as host, contorts his body around to get a beer, but maintains his upper-body orientation to Mike, and in this displays his audienceship. Meanwhile, Mike is fighting to be heard, his counter of 'No. Somebody

3.3

```
Mike:   =·hhhhh So they all┌go dow┌n
Gary:                      └A:ll  └All show
             (0.2)
Carney: Yeah, th┌ey all,=
Mike:           └They all-
Gary:   =hn-┌hn!
Mike:       └They all go down th┌ere,=
Gary:                           └°Gimmie
        a┌beer Curt,
Mike:    └=N┌o some-somebuddy so:mebuddy,
Carney:     └It reminds me of those
        wrestl(h)ers. 'hhh
Mike:   So:me┌body ra:pped=
Carney:      └hhh(h)on t(h)elevi┌sion. °(         )
Gary:                           └=Bartender how
        about a beer. While yer
        settin┌there.
Carney:       └°(      ).
Mike:   So:mebuddy rapped uh:.
Curt:   °((clears throat))
Mike:   DeWald'nna mouth.
Curt:   Well, h┌e deserved it.
Mike           └But yihknow eh- uh-he mae iz
        first mistake number one by messin with
        Keegan because a'pits'r fulla Keegans
```

rapped Dewald in the mouth' taking several attempts to produce in the clear, and him generally trying to move the telling forward to further events that are consistent with his 'fight' version.

By the end of the story, Mike's audience is different and more limited than the one that listened to the story's opening. Early on, one of the somewhat excluded recipients is not only able to offer a way of understanding the events that the main teller is describing that undercuts the seriousness and drama the teller attributes to them, but to also recruit others to that position. This changes the structure of the interaction, making byplay more permissive, and facilitating Gary's later alternative climax.

Empirically, the neat classification of byplay is challenged here, as while this listener to listener alignment fits the general idea of interaction between ratified listeners, it's arguable whether there is actually subordinated conversation. More generally, speakers propose a certain alignment to the story being told, and this is done through their gestures, intonation, word selection and arrangement of events. Likewise, listeners have resources available to them for analysing the talk that is being heard, aligning themselves to it in a particular way, and participating in the field of action it creates. The ensuing interaction can be with the speaker, as per the discussion of addresses in the previous section, or within the audience –

or, as here, both.

This is a kind of performance, and both the quality of narration and the listeners' experience has been materially affected by interaction within the group. Goodwin (1986) shows how storytelling provides an arena within which participants can test, negotiate and establish their competence and standing in relation to each other. Any heterogeneity in the audience challenges the speaker and audience members alike to control the emerging narrative. The result is that the audience can be seen to be both shaped by the talk it is attending and help shape what will be made of that talk.

3.1.3 Mass-interaction

Dialogue necessarily orients the analysis around a single person, the speaker. How might multi-party human interaction look without that orientation?

Asking what makes a crowd a crowd, Reeves, Sherwood and Brown (2010) explored the material practices of sports fans as they interact and participate with each other. The central insight is that crowds are 'see-able', a quality used by those present to navigate the situation, and available for analytic observation. Two vignettes derived from video recording of pre-match activity of football fans in a pub emphasise that when in this environment you see and hear seemingly unlinked groups of supporters doing a number of different things in concert with each other. Considering each of these things, such as singing, jumping, gesturing, striking a plastic deer, they find a consistent orientation to sharing and coordination. For example, the song sung carries and has a repeating structure, so is audible to all present in the venue and presents the opportunity of joining in on a repeat. Jumping and raised fist-pumps are visible beyond the immediate group, and present a similar opportunity of being performed together. Taking part in a strike of the deer involves participation between members of the crowd in distinctly different spatial places; the deer may be seen as a kind of 'shared object' for members of the crowd to coordinate with and around. Overall, Reeves et al. demonstrate a sensitivity in the production of fan's own bodily and verbal conduct to those immediately around them and, crucially, to those further away. The suggestion is of *crowd-ness* as a sort of distributed performativity requiring pragmatic skills and understandings.

In this view, a crowd becomes manifest through collective participation. This is plainly apparent when people are doing things together: an assembly of people is clearly a crowd when those people are chanting in unison. This is also apparent, if more subtle, in the conditions that lead to such collective acts, where people design their actions to the possibility of collective participation. Crowds can then be understood as where such design is pervasive.

Further, the collective participation of a crowd is not homogenous, but based on fluid groupings of individuals. Crowds are possible where members hold membership of such

groupings open, providing a participatory structure for individuals to adopt the behaviours of others. Emergent behaviours can then be seen to ebb and flow across assemblies of people, while at any moment being the specific actions of mutually identifiable individuals.

The joint project and structured audiences of dialogue are based upon the possibility of mutual monitoring. Here too, such crowds are based upon the possibility of mutual monitoring. The critical difference is that the structures of participation are no longer related to a central speaker, but rather in terms of distributed, fluid groupings; a network analysis, in effect. At a local level, Goffman's notion of ventures in joint orientation holds (section 3.1.2, *Beyond the dyad*), even if jumping up and down in tandem, or an ephemeral alliance around a giant inflatable banana, isn't necessarily what he had in mind.

3.1.4 Actions in the presence of others

Goffman's (1959) observation that life is a performance is well established. The role of gaze in dialogue has already shown we look not just to see, but to be seen looking. Sensitivity in the production of a fan's bodily and verbal conduct to those around them made 'crowd-ness' possible. How far might this go?

One approach is to see whether a behaviour is only present when there is somebody to direct it to. Studies have demonstrated smiling as communicative behaviour using this approach. Observing bowlers, Kraut and Johnston (1979) report that of 116 bowls made by 34 people smilies were seen 36 times while the person faced the rest of the group versus four when not. Coding 1793 four second sequences after a bowl for various behaviours and whether it was a good score, a temporal co-occurrence analysis of smiling with other behaviours and a principal component analysis of similarities of co-occurrence both suggest that bowlers smiled when they were being social, when they were being playful, or when they were otherwise communicating an emotional statement to an audience. They report similar findings obtained by coding photographs of ice hockey stands full of fans, with the change of smiling when in a social group being 23% vs 6% when just facing the game. Although here something positive for the fan's team just having happened also increased the chance of smiling (13% vs 4%). These photos were taken immediately following events that were favourable, neutral or unfavourable to those fans. 3726 faces were coded, and grouped into those identified in social units (with one member facing the others and not the game) and those not. In a final study presented on pedestrians, social interaction was found a much more powerful predictor of smiling than whether the weather was nice or not. In all these cases, a social motivation for smiling was compared to an emotional elicitor, and social involvement was found a major cause of smiling, independent of the emotional elicitors.

Correlating presumed emotional elicitors with other phenomena does not mean the emotional state of all the people being compared is consistent, or even known. Fernández-Dols and Ruiz-Belda (1995) addressed this by conducting a study of facial expressions at one

of the happiest times of an athlete's life: being awarded an Olympic gold medal. Observers, including other Olympic gold medalists, judged the emotional experience of the athletes to be intense happiness throughout the ceremony. As the ceremony included two stages that restricted any social interaction surrounding a section where the athletes interacted with the authorities and public, a comparison could be made. This study, too, found happiness not sufficient for smiling, but an interactional situation makes smiling much more likely.

Even when alone, smiling has been shown to be better predicted by social context than reported emotion. Experimentally manipulating the presence of a friend, Fridlund (1991) report smiling increased with the sociality of viewing but not with reported emotion. Sixty-four participants viewed a pleasant video variously alone, alone but with the belief that a friend nearby was otherwise engaged, alone but with the belief that a friend was viewing the same videotape in another room, or when a friend was present. The participants' facial behaviour was measured using electromyography on the muscles responsible for smiling. The participants' emotional state was self-reported before and after the viewing using forms to score the extent to which various emotional states were felt. For the alone and for the explicit audience conditions, the results agree with the studies above. The novel enquiry and result was that an implicit audience also had an effect. Equal 'smile' muscle activation occurred in the mere psychological presence of a co-viewer as with the direct presence of a co-viewer.

These studies show selective behaviour depending on social context; the smiles are a social display. Do people – audiences? – selectively display their *reactions*?

Bavelas et al. (1986) show motor mimicry – e.g., wincing when seeing someone hurt – is precisely controlled by the degree and timing of eye-contact. Testing further they find the mimicry is produced in the form it can be received (Bavelas, Black, Chovil, Lemery & Mullett, 1988). These ostensibly spontaneous reactions are shown instead skillfully inserted into interactive sequences. They are explicable as being *to* another, not *of* the person's feelings.

A scenario was staged where an observer saw a workplace accident. Two variants were choreographed, each holding the apparent injury constant while varying the likelihood of eye-contact between confederate and observer. Each four second choreography was designed to sustain a sufficiently fine-grained analysis to reveal any interactive sequence in addition to a broad determination of degree of response. The recordings showed that the presence and availability of a receiver affected both the pattern and timing of motor mimicry. In the eye-contact condition, 14 out of 15 participants displayed some expression of pain, and as the probability of eye contact with the receiver increased, motor mimicry not only increased generally but was available at the best opportunity for its receipt. In the no eye-contact condition, 12 out of 17 displayed some expression of pain, and for those that did, it faded away quickly as the probability that the receiver could see the expression diminished.

In subsequent interviews participants from both conditions produced mimetic expressions indistinguishably when prompted to recall what they saw, i.e. without the stimulus but with an available receiver. The authors' analysis states motor mimicry as no more and no less than the overt behaviour visible to others, and then asks what other behaviours, processes or constructs is the phenomenon related to most directly. The differential effect of visual availability of a decoder requires at least part of the explanation to be, somehow, between the individuals, i.e. social. The precise synchrony (tenths of a second) of the motor mimicry to visual availability rather than the accident onset suggests social context to be not just some part, but the primary factor. A social view also fits the interview data; in both cases, whether when injury was taking place or later when it was being described, the participant was representing a reaction and conveying this representation.

For this representation to be communicative, the produced representation would then be meaningful to others. This was tested by having participants 'decode' and rate excerpts of the recordings. If the mimetic expressions had no message value to the decoders, these results would be random. Instead, the evidence is that the behaviours of participants in the presence of a visually available victim produced a more knowing and caring message than behaviours in the presence of an unavailable victim. Together with the visually available receiver evidence, the authors demonstrate these were not just informative acts, but communicative ones: iconically encoded, inserted skilfully into the interactive sequence, and consistently decoded by receivers.

Throughout this section on the dynamics of interaction, mutual monitoring has been the constant factor. Through mutual monitoring, the collaborative, simultaneous producing and receiving of dialogue is possible. Through mutual monitoring, the emergent behaviours of crowds are possible. In this section, mutual monitoring greatly expanded what might naively be thought of as a social display; between social smiles and the empirical fact that audiences selectively display their responses a number of hypotheses about the audiences of stand-up comedy become possible.

3.2 Performance?

Having explored mechanisms of interaction and their consequences, might this apply to live music, comedy and drama? Performance emphasises the performer, this imbalance does not fit with the broad, peer-to-peer symmetries of dialogue. However, asymmetric dialogue has been studied, along with concerns about the collaborative model suiting only task-based dialogues where both parties have information to share. Even in highly asymmetrical dialogues, speaker and listener roles are not fixed and separate. Rather, their relationship is reciprocal and collaborative, in that the speaker elicits responses from the listener and the listener's responses affect the speaker. This has been experimentally shown, and shown for a particular form of performance that no matter how good the plot is to a story you're

telling, a good listener is crucial to telling it well (Bavelas, Coates & Johnson, 2000).

A narrator telling a story to a stranger may at first appear to be the delivery of a polished monologue to a passive listener, but instead Bavelas et al. show in a series of one-on-one, co-present, 'close-call' storytelling sessions that the listener responds to the narrator, and the nature of their multi-modal feedback had demonstrable effect on the quality of narration. Examining closely what the listeners actually did, the authors identify generic responses such as nodding and "mhm" – typical *backchannel* responses, serving as a cue for the speaker to continue (Yngve, 1970) – and a further, *specific* kind. The generic responses are highly contextual, being timed precisely and appropriately to the narration. However these responses lack any narrative content. Specific responses are tightly connected to what the narrator is saying at the moment, such as looking sad, gasping in horror, mirroring the speaker's gesture, or supplying an appropriate phrase. As such, they expand, reinforce, embellish the narrative, making the listener a co-narrator in the story. Such co-narration was readily seen in the study's data: often the speaker only described the facts while it was the listener who through facial displays illustrated the dramatic or emotional import of the facts.

The final experiment ran as follows. In 24 unacquainted dyads a speaker told his or her own close-call story. In each dyad, who was the listener was randomly assigned, and this listener was randomly assigned to either listen to the speaker's narration or the speaker's words. The narrative condition tasked the listener to listen to the story such that they could summarise the main point of it to someone else. The words condition tasked the listener to press a (covert) button whenever a word was spoken by the narrator that contained the letter *t*. Dyads that did not demonstrate attention to the words, judged by the accuracy of their button presses, were discarded. In both conditions, the narrator knew only that the listener was going to be listening for something in the narrative.

These storytelling sessions were recorded and analysed. For each session, a face-on view of the speaker and listener was split-screened into a single audio-visual recording. A coding scheme was developed, defining two types of listener response:

> *'A' responses keep the listener clearly in the role of listener/audience. That is, they are made by the listener as audience or observer and are external to the narrative plot. [...]*
> *'B' responses are co-telling acts in which the listener becomes (however briefly) involved with the narrator in the telling of the story. The listener is more of an actor than an observer of the story. 'B' responses are internal to the narrative plot. [...]*

Two independent analysts identified generic ('A') and specific ('B') listener responses in the recordings, noting their time of each occurrence; checks established 95% overall agreement between the two. Story endings were then evaluated. Points were scored for negative features relating to the following questions: Was the pace of the ending appropriate or abrupt? Did the narrator appropriately extend the denouement or just talk on and on?

Was the ending choppy or not? Did the narrator try to justify or explain the closeness of the close call or not? Two analysts worked independently and then collaborated to produce a set of ratings, which were then compared to a third analyst's ratings for reliability (r=.69). Descriptive statistics of the quantitative results for the two conditions were produced, and compared using the t-test, a way of assessing whether the averages for two groups are statistically different. A further test of the results mapped the same analysis procedure to the recordings of an earlier iteration of the experiment with a less demanding narrative vs. words manipulation, resulting in an analysis of 63 dyads in total; these results were consistent with the final experiment results.

This experimental design allows for more compelling evidence than the basic co-narration observation above. Based on the prediction that if the listener is distracted from the dialogue they should be less able to make appropriate listening responses, the experimental conditions manipulate the listener's level of involvement versus distraction to the narrative. This was done while ostensibly holding the listeners' attention constant.

Following the conjecture that listeners do not or cannot make specific responses until they have enough information about the narrative, generic responses were identified throughout the narrative whereas specific responses occurred later in the narrative. Listeners who were experimentally distracted from the narrative still made some generic responses ($t_{(22)} = 0.78, p = .45; U = 41; p = .034$)[2] but almost no specific responses ($t_{(22)} = 5.94, p < .0001$); the word-counting task reduced the rate of generic responses to 80% of their rate in the narrative condition, whereas specific responses dropped to less than 5% of the rate in the narrative condition. There was a substantial difference in effect sizes; for specific responses $\eta^2 = .62$, for generic responses $\eta^2 = .03$. Most striking is that the experimental conditions that distracted the listener from the narrative also strongly affected the narration itself. The story ending analyses showed that the stories faltered or fell fiat when they were told to listeners who were attending closely to the individual words but not to the narrative itself ($t_{(20)} = 4.25, p < .0001$).

The responses were cued by mutual gaze (Bavelas, Coates & Johnson, 2002). An inductive exploration of the storytelling data for anything that had happened just before a listener response had led to mutual gaze being identified for further analysis. The timings of onset and offsets of mutual gaze were then taken alongside the existing data for onset timings of listener responses, for nine dyads. Statistical tests confirmed that for each dyad placement of listener responses within periods of mutual gaze were not simply a coincidence: the listener tended to respond when the speaker looked at them, and the speaker tended to look away soon after the listener responded (one-tailed binomial test gave $p < .002, .009, .0001, 0.05, .0001, .0000, .0003, .0000, .008$). Together, speakers and listeners created and used this *gaze window* to coordinate their actions. Further tests

[2]Generic response results were re-evaluated using the non-parametric Mann-Whitney u-test, to account for a few outliers who managed to make a high number of generic responses in the word condition

explored the actual presence and visual availability of the story-teller, by having stories told through a recording, via telephone, separated by partition, and fully face-to-face (Chovil, 1991). Analysing facial displays of motor mimicry, the results show these displays are mediated by the extent to which individuals can fully interact in communicative situations.

The listener had to be following the meaning of the story closely to be able to insert precise and specific responses at appropriate points in the narrative. This shows that a good listener must be not merely attending but attending at the level of the narrative, and that the responses are best understood not in terms of a lexicon or bounded by individual modalities, but as multi-modal signals precisely configured, moment-by-moment, to the evolving narrative.

Bavelas et al. show a relationship between i. an audience member's (in)attention and the nature of feedback they give, and ii. the nature of such feedback and a performance's on-going quality. Their qualitative analysis shows a process of collaboration on the task at hand – advancing the narrative – and the quantitative analysis shows a 'micro' process coordinating that collaboration – with timings resolved to tenths of a second. Through one carefully designed and iterated experimental task, the authors provide a cogent account of how interaction is key to both the content of these performances and its delivery.

Successful one-on-one storytelling requires that the listener act as an addressee in dialogue. In the terms of section 3.1.1, *Dialogue*, the actions of an addressee are transformative. Section 3.1.2, *Beyond the dyad*, observed similar deleterious effects on the storyteller's narration in a small group situation, suggesting the actions of a by-player are similarly transformative. This is strong evidence that interaction is highly relevant to the study of performance. What remains an open question is the massed case of section 3.1.3, *Mass-interaction*.

3.3 Audiences?

Specific kinds of performer–audience interaction have been identified, but an audience of one is not the focus of the present work. Is there evidence that such highly local, fine-grained interaction is a feature of the large audiences found in most live events?

One body of evidence concerns public speaking. Collective audience responses such as applause and cheering have been shown to be closely synchronised with the speaker's actions. Certain rhetorical devices used by the speaker in set-piece political speeches were found to correspond with audience applause by (Atkinson, 1984), who had the insight that the structure of these *claptraps* allowed the audience to anticipate when a response was desired, and so collectively provide it on cue. Elaborating from speech extracts to a comprehensive review, Heritage and Greatbatch (1986) identified seven devices that corresponded with 68% of applause incidents across all of the 476 speeches of the British Conservative, Labour and Liberal party conferences in 1981. Gestures have since been suggested to signal whether

the rhetorical device is to be taken as an invitation to respond, with Bull and Wells (2002) accounting for more response incidents / non-incidents this way. In these accounts of *invited* applause (Bull & Noordhuizen, 2000), some of the responses identified in the previous section, so transformative to the quality of the performance there, can be seen to made collectively, at scale.

Uninvited responses emerge quite differently. Studying the occurrence of booing in various US and UK contexts, Clayman (1993) observed that collective booing is often preceded by other audience behaviour. This may be minor displays of disaffiliation such as audience members whispering or talking among themselves, groaning, shouting or jeering at the speaker. The resulting "murmur", "buzz", or "roar" then would transition into the booing. But booing was just as often observed to follow affiliative displays such as applause or laughter, in this case appearing as a competitive counter to an initial favourable response made by others. In neither case was there a suitably timed speaker action to cue the response. Rather the individual response decisions appear to be made, at least in part, by reference to other audience members. In this account, audience responses with the specificity emphasised in the previous section are being produced collectively. However, the audience behaviour is no longer explicable as merely an aggregate of independent responses cued by a single stimulus.

The audience behaviour is however explicable by considering general interactional practices that condition acts of affiliation and disaffiliation across diverse contexts. Applying the prompt display of affiliative responses seen elsewhere (eg., preference organisation – Sacks 1987; Pomerantz 1984) to audiences imposes a timing constraint only independent decision making can uphold – there is no time to check for others. The preference to delay disaffiliate responses however offers that time to check for others and the opportunity of other responses to act as triggers. Clayman thus distinguishes between coordination as an aggregate of independent responses, as per the discussion of invited responses above, and coordination amongst an audience through a process of mutual monitoring.

What is less clear, is in the detail of how a collective response is achieved. A game theory approach is suggested by Clayman for booing, where the incipient actions of the other audience members reconfigure the costs and benefits of initiating a booing response in a variety of ways. One is fear of being isolated, whether being exposed as the only person making response, or the only person not. Another is making a message explicit, in some cases transforming essentially disparate private displays into an unambiguous, collective public display, in others transforming a unified collective response into a more representative split response.

Both Atkinson (1984) and Clayman (1993) theorise public speaking as a kind of dialogue where speaker and audience take conversational turns. The timing and repertoire of collective audience responses appears more limited than seen in one-on-one performance,

where Bavelas et al. (2000) emphasise a repertoire of simultaneous but non-interruptive responses enabled by being co-present: nods, smiles, confused looks, winces, and so on. Nonetheless, the broad narrative is similar: that something that at first might appear as monologue is better understood as a dialogue between performer and audience.

A study of theatre audiences shows byplay-like interaction, rather than the addressee-like interaction of public speakers and their audiences. In a video-ethnography of a number of performances of contemporary French plays, Broth (2011) makes the basic observation that the audience is mostly quiet, and that they act in particular ways when unable to be completely quiet; there is a norm of silence to be upheld, and evidence that the audience do this. Vocal noises such as coughs and throat-clearing tend to appear within dramatic actions as they finish, leading to an overlap that partially masks the vocal noise, or when the performance is 'off' such as between scenes. Laughter appears between dramatic actions, avoiding any overlap, and is only made aloud with others, such that the individual act is lost in the collective response. Further, during audience laughter, actors may 'stretch' the performance by the use of gestures in order to enable the production of the collective laughter before the subsequent dramatic action. Such phenomena are more striking considering that it has been experimentally shown that the larger the audience, the more coughs occur per person (Pennebaker, 2010) (see also Wagener 2012).

In the theatre, Broth argues, there is conversation-like turn taking, but it is between the interacting actors on stage. As it is the script that selects whose turn is next, a turn by the audience will never be due. Nonetheless the actions of the theatre audience are accountable in terms of dialogue, in the ways discussed in 3.1.2, *Beyond the dyad*. They place their vocalisations carefully in relation to this on-stage turn-taking, working constantly to be able to do so. Vocal noises are subordinated. Laughter, however, can constitute a turn. Inserting laughter is done by achieving collective laughter shortly after a possible ending of an on-stage turn, and exiting is done by stopping or minimising laughing before the actors take the turn back, which seems to be mainly achieved by exploiting the projectable properties of actors' body movements. Notably, the subordination allows the fictional world being presented to be unaffected, despite the material impact on the performance as audience and actors interact to allow the turn.

Highly local, fine-grained interaction is thus shown to be a feature of these ostensibly passive theatrical audiences. The ways in which both vocal noise and laughter are produced can be seen as public displays of close and submissive attention, managed as part of a sequential, interactive process.

These ethnographic accounts show that the highly local, fine-grained interaction found transformative in one-on-one performance is a feature of the large audiences found in at least some live events. Collective responses are achieved and inserted skilfully to the sequence of action on stage. The sequential detail of these collective responses suggest

pervasive mutual monitoring: between performer and collective audience, and amongst the audience. However, these studies do not provide evidence at the level of the individual, and lack the kind of contrastive rigour shown in the previous section on performance. Making an argument about audiences in general will require better evidence.

3.4 Motivating an experimental method

In live events, interaction matters. And interaction is not an abstract thing. This chapter has shown specific interactional mechanisms that are well understood in everyday contexts are relevant to the study of live events. The suggestion is that live events are to some degree social. There is the possibility of mutual monitoring, and evidence that this is being acted on.

Most of all, this chapter has elaborated on the active nature of spectatorship; that we are actively engaged with each other and performer, reciprocally. So is there a distinction between people who are merely massed together and *audiences*? Yes, and it is in the kinds of social organisation involved. The kinds of collective audience behaviours, reported first in street performance (section 1.2.2) and stand-up comedy (section 2.1), and elaborated here are only possible when people produce their own actions with a sensitivity to those around them.

Are the kinds of social organisation distinct to audiences well understood? Not really. Where the ethnographic accounts of section 3.3, *Audiences?*, identify intra-audience effects, they do so from evidence that only captures overall audience behaviour, and often only audio. There may be some ecological validity to this, as for instance audiences are often in the dark, while sounds can carry across the auditorium; the ability of a researcher to identify a single audience member's cough in an audio recording made of the overall audience may well be a fair representation of what is available to an audience member. Nonetheless, the discussion of mutual monitoring is necessarily speculative.

How then to understand these interactional dynamics in a sufficiently coherent way that they might be applied in a different genre of performance? Or to identify the mechanisms and processes at a sufficient level of detail to enable successful practical interventions in such events? Reviewing the chapter, the most systematic evidence has come from experimental contrasts; in particular the contribution made by the one-on-one storytelling study that informed the performance section. How might such an approach be made for audiences? There are some precedents.

Audience effects have been experimentally shown for small audience sizes. During public speaking, an audience giving negative feedback provokes an anxiety response irrespective of the normal level of public speaking confidence of the subject (Slater, Pertaub & Steed, 1999; Pertaub, Slater & Barker, 2002). This was found when presenting in a virtual reality environment to an audience of eight avatars, taking advantage of the potential

for experimental control of such an environment. Audience responses were largely non-verbal: for an audience giving positive feedback, avatars nodded encouragingly, smiled frequently, leaned forward and orientated their bodies to face the speaker; for an audience giving negative feedback, avatars falling asleep, slouched in their chairs, slumped forwards on the seminar table, oriented themselves away from the speaker, leaned backwards, put their feet on the table, avoided establishing eye contact and one even got up and walked out of the seminar room during the talk. Similar audience responses have been used in virtual reality classrooms, testing effects between listeners. The experimenters were able to pare back the features of the classroom environment and participants, while experimentally manipulating configuration and behaviour – often in ways impossible in a real life, for instance where the rendering of the teacher in each student's view has the teacher looking directly at the student, for each student simultaneously. Students were found to conform to the behaviour and learning abilities of student avatars around them, and manipulating eye gaze and distance from a teacher to student affected the student's recall of the lesson topic (Bailenson et al., 2008). In these studies, the audience behaviour surrounding the human participant is autonomous, and so these are not interactional situations in quite the sense developed so far. They do, nonetheless, highlight the importance of monitoring others in these situations: all the effects demonstrated are attributable to monitoring others in the environment.

Also relevant is the work of Mann, Faria, Sumpter and Krause (2013), which *is* predicated on analysing individuals in the audience rather than the audience as a whole. Recall that Clayman's independent and mutual monitoring coordination processes were inferred from gross audience characteristics; a sudden burst versus a staggered build (section 3.3, *Audiences?*). Simulations of such characteristics have since been compared to empirical results, quantifying the role of mutual monitoring in the start and stopping of applause. Drawing from studies of contagion in disease and flocking behaviour, Mann et al. (2013) mathematically modelled applause spreading independently of others, in proportion to the total already applauding, by influence of local neighbours, or after some tipping point. Bayesian model selection was then used to test these hypotheses against experimental data, where 107 students in groups of 13-20 listened to two presentations each, providing data for the times at which each individual started and stopped clapping. Both the onset and cessation of group clapping were found to follow a sigmoidal curve, with an initially slow uptake of the new behaviour followed by a phase of rapid change and eventual saturation. The analysis showed that once the first person had started clapping, the probability for another individual to start is proportional to the total already applauding, regardless of their spatial proximity. Such linear 'social contagion' was also found for stopping, although here the analysis could also include the non-social factor of duration of clap (as start to stop is provided in the data, whereas presentation end to clap start could not be due

each presentation's different and ambiguous ending). This analysis found 'how long have I clapped' was a factor, although secondary to the 'how many others are still clapping' factor. Consistent differences were also found between individuals in their willingness to start and stop clapping.

Chapter 4

Experimenting with performance

In chapter two, the compère of a comedy club was observed and analysed. Here, the empirical focus moves to a stand-up comedy act. A live performance experiment is staged that tests audience responses to a performer's gaze and gesture. This chapter provides the first direct evidence of individual performer–audience dynamics within an audience, and establishes the viability of live performance experiments.

4.1 Methodological issues

Experimenting with live performance and audiences is a departure from the tasks and dyads of much of the literature reviewed in chapter 3, *Audiences and interaction*. This presents a number of methodological issues.

How might a live performance be systematically manipulated? The performer–audience dynamics that need to be probed are between performer and audience individuals; chapter three shows that there has been some study of collective behaviours in the audience, but there is paucity of evidence at the level of the individual. Thus, a performer needs to do something directly at an individual. To become representative of the audience overall, this could not just be between the performer and one individual, but would need to be repeated across as many individuals as possible. This could actually be quite simple, for instance just looking at people in the audience would provide a contrast, allowing their behaviour to be compared before, during and after the fixation. The methodological issue is how to do this while maintaining experimental control. Could, for instance, a comedian be tasked with looking at everybody in the audience? With some kind of random sampling and whatever other controls good experimental design might necessitate? This would require considerable control by the comedian of their behaviour. It also might compromise their performance, conflicting with the acts their dramatic instincts would otherwise be leading to. Intuitively, such a performance could get very weird, very quickly.

Such concern is borne out in the literature. There is evidence that the using confederates in dialogue experiments (e.g. as conversational partners) leads to different results than when the dialogue is between two freely conversing parties (Bavelas & Healing, 2013) . The basic case is that many of the phenomena that occur naturally in interactive conversation are difficult, if not impossible, to create in scripted utterances and within traditional experimental paradigms (Brown-Schmidt & Tanenhaus, 2008). Asking when a confederate can fulfil the role of conversational partner without changing the nature of the dialogue itself, Kuhlen and Brennan (2013) further elaborate the case and propose experimental manipulations be developed that enable naive partners to fill roles typically filled by confederates. One such manipulation has been seen already in the one-on-one storytelling of section 3.2, *Performance?*, where instead of asking a confederate to act inattentively, a naive addressee was secretly preoccupied with a second task.

It is likely there is no neat solution to this issue of experimental control of live performance. While the everyday skill of dialogue allows participants to be assigned the role of speaker, addressee, or otherwise, this is not so for these kinds of stage performances. If a confederate is required, then as per Kuhlen and Brennan (2013) they will need to be recognised as being active and potentially influential participants in the interaction, whose behaviour needs to be systematically understood, managed, modelled, and monitored.

Another issue is one of effect. How might a manipulation's effect on an audience be

known? Section 3.1.4, *Actions in the presence of others*, noted that given social smiles and peoples' selective display of responses a number of hypotheses about the audiences of stand-up comedy become possible. One could be that a person in the audience would smile back at a performer looking at them, a kind of humour-specific backchannel "mhm". So developing the posit, above, of the performer simply looking at an audience member vs. not, a well motivated measure would be to see if their facial expression changes when fixated upon. What is vital to note about this example, and the social organisation argument more generally, is that it is only *what is there for others to see* that matters (or hear, or...), and it needs to be analysed in the context of that moment. Note that all the studies in chapter three took such an approach. Without diminishing any notion of inner, private lives, they provided perspicuous accounts of human behaviour while studying only what is externally manifest.

This means many measures of audience response used elsewhere are inappropriate, most obviously questionnaires and suchlike that rely on subjective assessments. This also aligns with the suggestion made in section 1.1.4, *Liveness?*, that the dynamics of the in-situ experience of live performance are unlikely to be available for conscious report. Relatedly, anything that requires an overt act in the moment, such as making that assessment on some device, would mask or disrupt any social display (e.g. Stevens, Dean, Vincs and Schubert, 2014). Also inappropriate are physiological responses such as galvanic skin response (GSR; e.g. Wang, Wong et al., 2016), electroencephalography (EEG; e.g. Dmochowski et al., 2014; Zhang, Sherwin, Dmochowski, Sajda and Kender, 2014), salimetrics (e.g., Silva, Soares, Miranda, Pereira and Harper, 2013), and so on, as these responses are unavailable to others.

Taken together, it is not at all clear whether such experimentation with performance is viable, or effects will be found. This is exploratory work.

4.2 Instrumenting an audience

In the interaction focussed literature discussed in the previous chapter, the data typically come from audio-visual recordings and a manual process of annotation and transcription follows. While time-consuming, this is a reasonable proposition for analysis of small groups. As the group size increases however, so does the manual burden, and at some point the approach will falter. How then to instrument a whole audience?

Automatic measures were sought. Could the auditorium-audio approach typical of the studies in section 3.3, *Audiences?*, be refined to capture more individual behaviour? The work for section 2.1, *Observing stand-up comedy*, had shown that to a comic a few people laughing loudly was a very different situation than many people laughing less so. This offered the possibility of automatic classification from the stage, but audio-processing techniques sufficient to the task were not found. Recreating, in some sense, the performer's ears is an approach with ecological validity and technical promise but is beyond the scope of the

present work, a research project in itself. A brute force approach could be tried instead, with ever-more microphones used to increasingly isolate audio to audience individuals. This would be at the price of rapidly increasing complexity and intrusiveness, however. Instrumenting for audio cues was abandoned on technical grounds.

Visual techniques proved viable, however. A single high-definition camera could capture a massed audience. Faces within the video image could be found and tracked by off-the-shelf software, as could features or expressions of those faces. In particular, Fraunhofer SHORE™ computer vision analytics[1] could operate in real-time across many faces in a scene, and offered appealing measures for comedy purposes, such as facial display of 'happiness' – the degree of smiling, in effect. The full measures offered were as follows (Küblbeck & Ernst, 2006; Ernst, Ruf & Küblbeck, 2009), requiring a minimum face size in the image of 35x35px (J. Garbas, personal communication, 21 March 2013).

Face tracking A face in the image will be detected and tracked across subsequent frames. Each tracked sequence is assigned an ID value, and the bounding box of the face is returned in pixel coordinates. This capability was validated using the CMU+MIT Frontal Face Images Database[2], obtaining 91.5% recognition rate at a 1 in 10 miss rate (ie. confidence to classify is set so one in ten is a false positive). Note no feature analyser was trained on the data with which their performance was verified.

Eyes, nose, mouth Coordinates for these facial features in pixel coordinates. Also for eyes: "Open" or "Closed". Also for mouth: continuous measure of openness.

Gender One of "Male", "Female", or "Unknown". This capability was validated using the BioID Face Database[3], obtaining 94.3% recognition rate.

Age Estimated age, in years. This capability was validated using i. the FG-NET database[4], obtaining 6.85 mean deviation in years and ii. the FERET fafb database[5], obtaining 92.4% accuracy.

Facial expression Continuous measures of the degree of "Happy", "Sad", "Angry", and "Surprised". The happiness analyser was evaluated on the JAFFE database[6], obtaining 95.3% recognition rate. The other three are unreported.

Rotation Detection of up to 60° of face rotation.

An academic license was applied for and use of the demo application was granted. Tests were performed. One worry was that such software would not be able to process imagery

[1] For more information see the developer's website: http://www.iis.fraunhofer.de/
[2] CMU+MIT — http://vasc.ri.cmu.edu//idb/html/face/frontal_images/index.html
[3] BioID — https://www.bioid.com/About/BioID-Face-Database
[4] FG-NET — nolongeronline
[5] FERET — http://www.itl.nist.gov/iad/humanid/feret/feret_master.html
[6] JAFFE — http://www.kasrl.org/jaffe.html

that wasn't brightly lit. The lighting required for the captured imagery to be adequate for SHORE was found to be quite low, allowing a convincing 'dimming of the auditorium' effect. A view of the lighting state actually captured in one performance can be seen in Figure 4.5. Another worry was that the analytics would produce 'garbage in, garbage out' measures. SHORE only reports faces when sufficiently confident, and the confidence for each reported face at that moment is one of the associated measures. While there are no confidence measures reported for each individual per-face measure, egregious data are simply not output; the actual measures only appear when, presumably, the source is adequate for that particular analyser. Two analysers appeared less useful: the eyes, nose, mouth measures were not as robust to lowered lighting conditions; the rotation measure reported zero throughout.

The application analyses imagery in realtime, ingesting live video or playing back video files. The analysis runs independently of the source frame rate, accepting the current frame when ready, and ignoring any subsequent ones until it is next ready. The application's user interface presents a window on-screen that annotates the image with bounding boxes and feature data for each face. The analysis data are also written out, in real-time, to a log file. This log file would then be the raw data for subsequent analysis, or could be acted upon during the event[7]. The logging was partly configurable; timestamps were added, pixel measures were normalised, and measures deemed non-informative were removed.

```
Frame=8   TimeStamp=2013-Jul-02   16:32:46.396849   Id=0   Uptime=0.352   Score=52.600   Gender=   Age=34.000   Surprised=0.000   Tra
Frame=8   TimeStamp=2013-Jul-02   16:32:46.396849   Id=1   Uptime=0.352   Score=40.000   Gender=   Age=37.000   Surprised=0.000   Tra
Frame=8   TimeStamp=2013-Jul-02   16:32:46.396849   Id=2   Uptime=0.352   Score=26.400   Gender=   Age=49.000   Surprised=36.000  Tra
Frame=8   TimeStamp=2013-Jul-02   16:32:46.396849   Id=3   Uptime=0.352   Score=63.400   Gender=   Age=25.000   Surprised=0.000   Tra
Frame=8   TimeStamp=2013-Jul-02   16:32:46.396849   Id=4   Uptime=0.352   Score=39.600   Gender=   Age=28.000   Surprised=0.000   Tra
```

— *Excerpt of logging script output*

A tool was developed to process this raw output into a form suitable for analysis.[8] ShoreExporter is a *Python 2.7* program of c. 300 lines. It performs the following tasks: i. transforms the one-line-per-found-object structure of the SHORE text file into a one-line-per-time structure, meaning the known state at any given time is compiled from multiple lines into a single line; ii. it transforms the timestamps into elapsed time iii. it identifies the found objects as the corresponding audience member through their spatial position in the frame; iv. it reduces the reported measures to a desired subset; v. it encodes all measures and no-known-values as numbers for ease of subsequent processing. The desired subset of measures, audience positions, etc. are set by a configuration file.

Usage: `python ShoreExporter.py input.log configuration.json`

ShoreExporter: `https://github.com/tobyspark/ComedyLab/blob/master/ShoreExporter/ShoreExporter.py`

[7] Reading back out in realtime a file that is being written to in realtime is plainly a hack, but so be it.

[8] ShoreExporter was implemented by Katevas, c.f. the robot engineering work in the next section.

Example configuration file: `https://github.com/tobyspark/ComedyLab/blob/master/ShoreExporter/configuration_05.json`

TimeStamp,	Frame,	Person02Happy,	Person02MouthOpen,	Person03Happy,	Person03MouthOpen,	Person05Happy,	Person05MouthOpen,
0.0,	8,	-10,	-10,	-10,	-10,	27.778,	0.000,
0.096005,	9,	91.111,	30.667,	-10,	-10,	38.889,	0.000,
0.225013,	10,	100.000,	56.000,	-10,	-10,	38.889,	0.000,
0.331019,	11,	100.000,	20.000,	-10,	-10,	38.889,	0.000,

— *Excerpt of ShoreExporter output*

4.3 Comedy Lab: human vs. robot

Comedy Lab: human vs. robot was a live performance event where one performance was experimentally controlled and the audience was carefully monitored. The rationale of the experiment was to use a robotic performer to perform a comedy script while different aspects of the delivery were manipulated and live audience responses gathered for analysis. Using a robot to perform stand-up comedy is clearly a radical intervention. Nonetheless, it seems no more or less implausible than other routes for maintaining experimental control (see section 4.1, *Methodological issues*), and has the benefit that such control will be absolute. As implementers, we will know what the interaction is, because we programmed it.

The opportunity to use an embodied robot came through the host organisation of this research, the Media and Arts Technology Programme at Queen Mary University of London. A colleague there had the opportunity of an internship at the company responsible for RoboThespian™, a humanoid robot designed for interaction in public places. The company gave the brief of developing the robot's capabilities for more nuanced interaction than it was capable of at the time. Observing stand-up had shown us, at Media and Arts Technology, an ideal challenge to drive such an improvement: could the robot successfully tell a joke to an audience? A concept for *Comedy Lab: human vs. robot* was drawn up by Kleomenis Katevas, Patrick G.T. Healey and this dissertation's author, Matthew Tobias Harris, and the engineering work was taken on by Katevas for the internship. In brief, Katevas found the built-in sensing and animation capabilities of the robot insufficient, and developed a separate script-parsing system that could accept real-time instrumentation of the audience and output real-time RoboThespian speech and pose commands. The script provides a monologue to be delivered, along with various audience-dependent behaviours. These behaviours included responses to make in response to audience laughter, and cues to look or point at an audience member. That work, separate to this research, is reported in "Robot Stand-up: Engineering a Comic Performance" (Katevas et al., 2014).

Why did the script feature cues to look or point at an audience member? Observing stand-up comedy had shown specific use of gaze and pointing to identify people in the audience. The compère analysed in section 2.1.1, *Ten minutes of Barry Ferns*, makes jokes at people's expense, with the point gesture or subsequent gaze shift 'landing' the joke

Figure 4.1: *Promotional poster for* Comedy Lab: human vs. robot

(e.g., "This woman here <points> was actually doing a white-power salute which is a bold <changes gaze> bold move. [...] You look terrified about the idea that you did that" in appendix A.1. He also incorporates and re-incorporates individuals into the content of his act in this way (e.g., gaze fixation and pointing accompanying "There's you and him...", "um... there was one over there." in appendix A.3, *167–247s: "Go Red Sox!"*). More generally, gaze shifted in a pattern similar to that observed in conversation where speakers use gaze to elicit responses from their addressees (section 3.1.1, *Dialogue*), effectively scanning the audience throughout the act.

The pilot study presented here is a collaboration between Katevas, Healey and this dissertation's author, Harris. It has been reported in *Robot Comedy Lab: Experimenting with the Social Dynamics of Live Performance* (Katevas et al., 2015). To stage our event, we were granted use of a RoboThespian by its developer Engineered Arts, and we collaborated with 'Hack the Barbican', an open-call programme of events at the Barbican Centre, an arts venue in London.[9]

4.3.1 Design

Experimental control of a performance is achieved by using an embodied robot to hold the 'content' of a performance constant (e.g., the prosody, semantics and syntax of the jokes) while selectively manipulating aspects of delivery (e.g., body orientation, gaze, and gesture). The credibility of the robot's comic delivery is established by i. establishing whether the jokes elicited positive responses amongst the audience and ii. comparing overall audience response to the prior acts, who were professional comics. This is achieved through automatic measures of facial expression, as described in section 4.2, *Instrumenting an audience*.

A within-subjects design is used to test two predictions. First, when looked at by the performer, audience members will display increased positive affect. By identifying audience members and dynamically selecting individuals as gaze targets throughout the performance, the conditions of 'just before', 'during', 'just after' are created. Second, when pointed at by the performer, members will display increased positive affect. By identifying audience members and dynamically selecting individuals as point targets at specific points in the script, conditions of 'just before', 'during', 'just after' are created.

These manipulations test whether displays of positive affect are conditioned by the social context – if any – of a live performance. The hypothesis is that audience members should display more positive affect when they believe the robot is attending to them and less when they believe it is not – independent of how funny they find the jokes, as suggested by the literature of chapter 3, *Audiences and interaction*. However, one obvious objection is that this prediction assumes that people will consider the robot part of that social context;

[9] "Europe's largest multi-arts and conference venue", "the Barbican pushes the boundaries of all major art forms including dance, film, music, theatre and visual arts." — http://www.barbican.org.uk/about-barbican/history, http://www.barbican.org.uk/about-barbican

after all, would you smile at a robot? If this were a factor, it would reduce any effect seen and so makes any evidence for the hypothesis more compelling. Further, there is precedent of using embodied robots in social science research (e.g., MacDorman and Ishiguro, 2006), and there are studies which show such robots being treated as social agents in the kinds of close encounters discussed in the previous chapter (e.g., Mutlu, Forlizzi and Hodgins, 2006).

4.3.2 Materials

On stage, a humanoid robot – RoboThespian – was controlled by bespoke *ComedyParser* software (see Katevas et al. 2014). Audience measures were obtained via an inconspicuous Gig-E Vision camera positioned high at the back of the stage with a field of view that encompassed the seated area. SHORE software then analysed video imagery from the camera in realtime. The resulting feed of face tracking and sentiment analysis data were passed to Comedy Parser. Figure 4.2 outlines the spatial configuration of the event, and Figure 4.3 shows the event mid-performance.

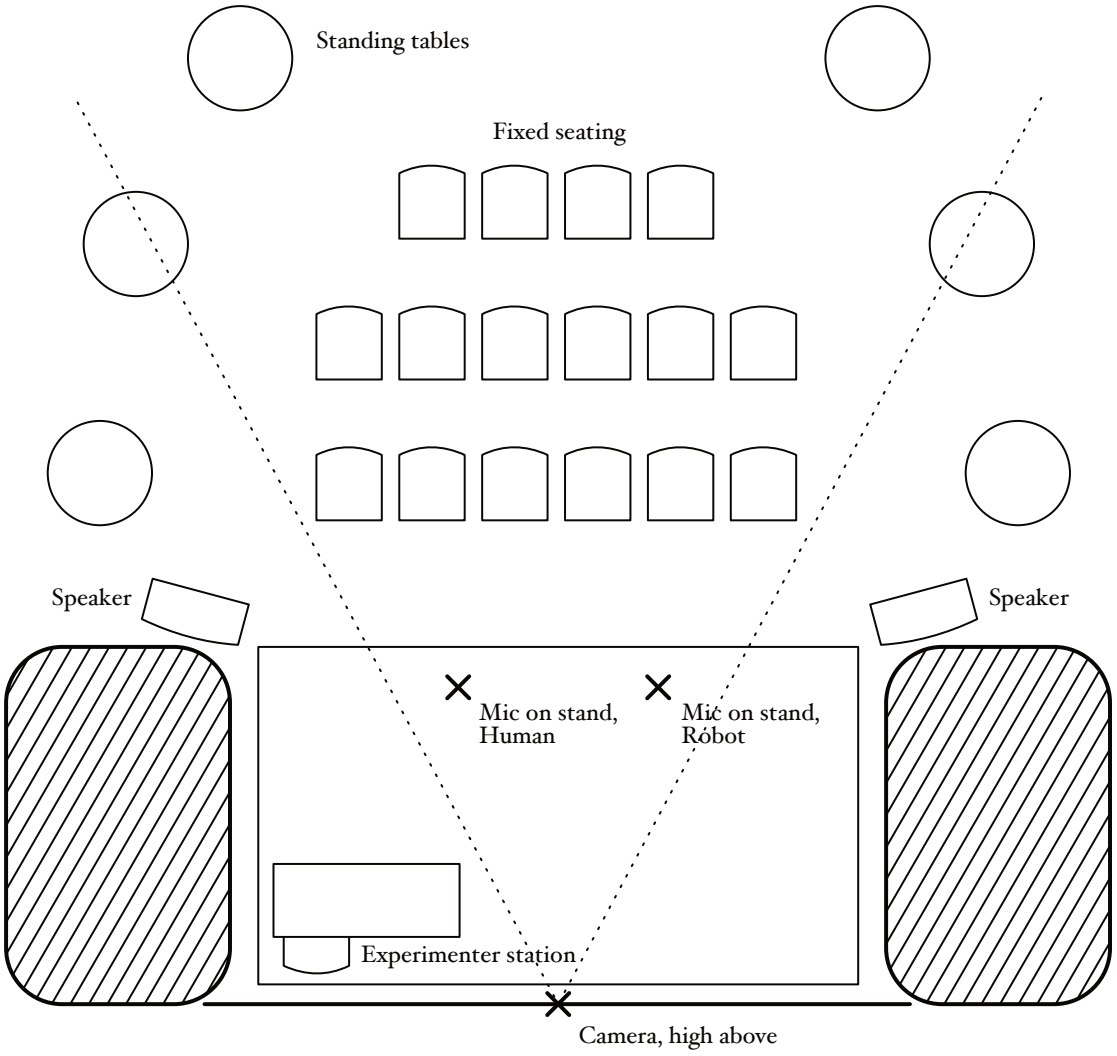


Figure 4.2: *Staging diagram for ComedyLab: human vs. robot at the Barbican Club Stage*

Figure 4.3: *ComedyLab: Human vs Robot, mid-show. RoboThespian™ is performing.*

A comedy routine for the robot was written by professional comedian Tiernan Douieb and then developed into a full ComedyParser script. The script encoded the experimental manipulation of gaze by commanding the random selection of a new gaze target from the audience at, broadly, every sentence end. The script commanded a "point" gesture at a randomly selected audience member accompanying the line "You go first", which followed "If you really want to be like us, here is a way to find your robot name. Take the first 16 digits of your credit card, and combine with the start, expiry dates and security code."

ComedyParser archived all dynamic aspects of the performance, in particular the robot's gaze and point. The SHORE data were also archived. Additionally, an audio-visual recording of each performance was obtained by placing an HD video camera toward the back of the audience area.

4.3.3 Participants

Audience participants were recruited by advertising *Comedy Lab* through social media channels of the two performers, the venue (The Barbican Centre, London), research group (Cognitive Science, Queen Mary University of London) and Hack the Barbican. The following context was provided in the advert:

> *What makes a good performance? By pitting stand-up comics Tiernan Douieb and Andrew O'Neill against a life size robot in a battle for laughs, researchers at Queen Mary University of London hope to find out more – and are inviting you along.*

A collaboration between the labs of Queen Mary's Cognitive Science Research Group, RoboThespian's creators Engineered Arts, and the open-access spaces of Hack The Barbican, the researchers are staging a stand-up gig where the headline act is a robot as a live experiment into performer-audience interaction.

This research is part of work on audience interaction being pioneered by the Cognitive Science Group. It is looking at the ways in which performers and audiences interact with each other and how this affects the experience of "liveness". The experiment with RoboThespian is testing ideas about how comedians deliver their material to maximize comic effect.

Approximately 50 people attended each performance on each night. Data from SHORE were captured for 22 people for the first night (15 men and 7 women between the ages of 28 and 64 years, $M = 46.4, SD = 8.0$) and 19 for the second night (13 men and 6 women between the ages of 27 and 60 years, $M = 46.2, SD = 8.1$). The discrepancy in numbers successfully instrumented is due to technical issues, c.f. section 5.1, *Instrumenting an audience, redux*.

4.3.4 Procedure

Two performances were staged as part of the Hack the Barbican event at 6 p.m. on 7th and 8th of August 2013 at the Barbican Centre. 'The club stage' was used, a venue within the complex which is freely accessible to the public. Each performance comprised the compère's warm up, and the (human) comedian's act followed by the robot act. The compère's warm-up lasted approximately 10 min, the comedian's lasted 13 min, and the robot's 8 min. Two professional stand-up comedians, Tiernan Douieb and Andrew O'Neill, were recruited for the compère and comedian roles, respectively. In addition to the credibility comparison discussed in the design section above, this format aimed to widen the appeal of the event and to help create a more convincing stand-up comedy context. Although the compère and comedian made normal stage entries and exits the robot cannot walk and therefore its position and the control desk were fixed throughout. During the robot's performance, an experimenter monitored the control equipment, visible to the side at the rear of the stage.

Participants were informed that they were being captured on video for research purposes and all data capturing and handling procedures were audited by the Queen Mary University of London research ethics committee (Reference: QMREC1199b).

The arrangement of seating, stage and robot performer can be seen in Figures 4.2 and 4.3, above. Figure 4.4 shows RoboThespian mid-performance. Figure 4.5 shows the view from just off-stage during the show, showing the experimenter's station, audience computer vision analytics on-screen, and the robot mid-performance. A summary video can be viewed at `http://tobyz.net/projects/comedylab`.

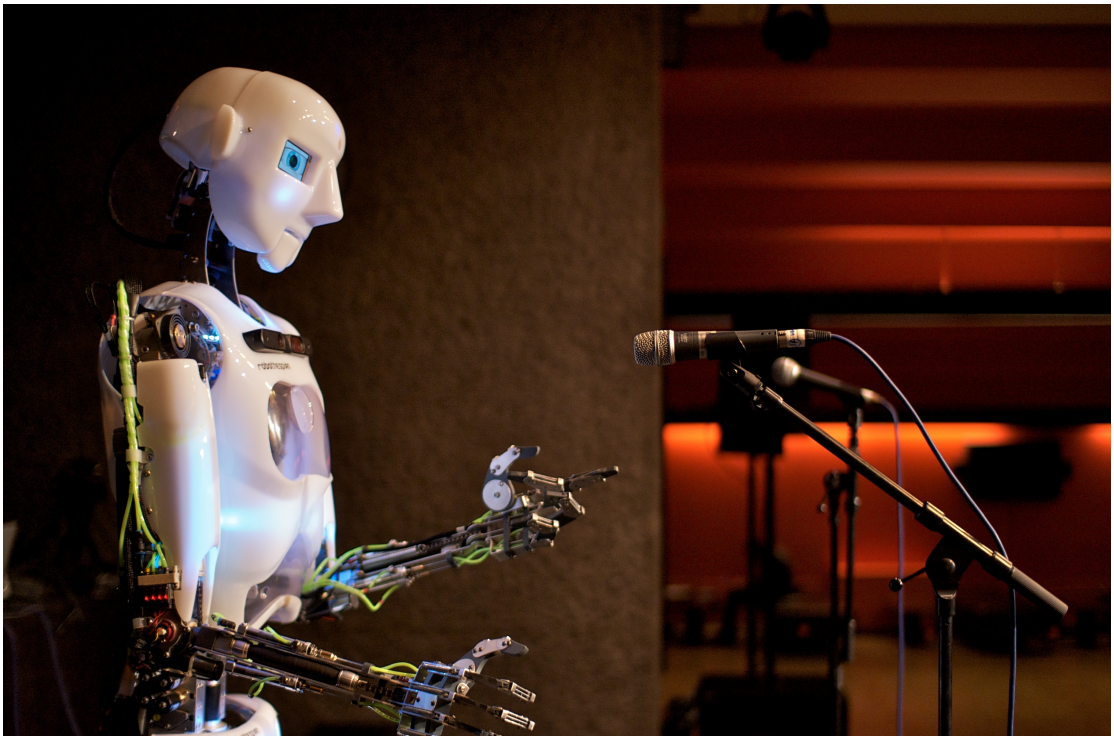


Figure 4.4: *A close-up of RoboThespian™, mid-performance*

4.3.5 Results

The measures of *Happiness*, *Anger*, *Surprise*, and *Sadness* produced by SHORE showed substantial inter-correlations. For example, in the data *Happiness* and *Sadness* are negatively correlated for: Pearson's $r_{(121)} = -0.484, p < 0.01$ (Note $N = 121$ because there are three measures for each of 48 people corresponding to *Before*, *During* and *After* a punchline, discussed in more detail below); and so are *Happiness* and *Anger*: Pearson's $r_{(121)} = -0.433, p < 0.01$. These correlations make these measures partially redundant and therefore results are reported only for the *Happiness* measure in the following analysis.

Throughout computed probabilities are reported for completeness. A criterion level of $p < 0.05$ is adopted for inferences. Generalised Linear Mixed Model (GLMM) analyses are used to model the combined random effects, categorical and interval fixed effects and repeated measures involved in the audience responses obtained.

The statistical analysis reported here is as per Katevas et al. (2015), and is the original work of Healey.

Credible performance: punchlines

To test if audience members respond selectively to the jokes, their facial displays of *Happiness* were averaged over three *ResponsePhases*: *Before*, *During* and *After* defined as 2 seconds before the punchline, the duration of the punchline delivery, and 2 seconds after. Punchlines were defined in the script (see appendix B, *Comedy Lab: robot script*).

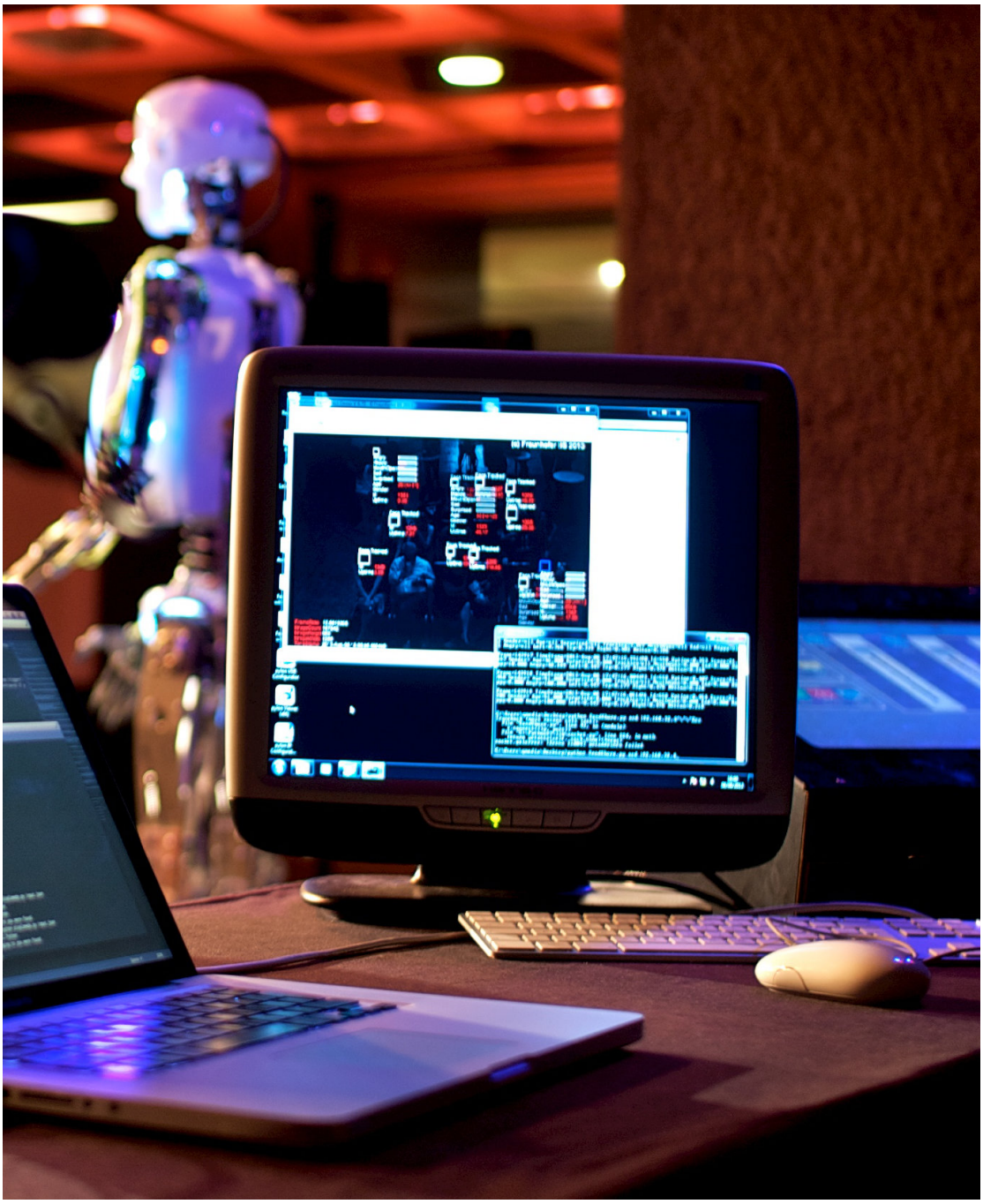


Figure 4.5: *View of live performance showing computer vision software extracting facial features.*

Average *Happiness* displayed by the audience was analysed in a GLMM using a Linear Model. This treated *ResponsePhase* (*Before* / *During* / *After*) as a fixed factor and *AudienceMember* nested within *Night* as random factors. It shows a main effect of *ResponsePhase* ($F_{(2,120)} = 5.66, p < 0.01$). Planned, pairwise comparisons show that people displayed more happiness after the punchlines than before them ($t_{(120)} = 3.32, p < 0.01$) or during them ($t_{(120)} = 2.67, p = 0.01$) but no difference in displayed happiness before and during the punchlines ($t_{(120)} = -0.86, p = 0.39$). The estimated means and standard errors are summarised in Table 4.1. Fixed (B) coefficients provide estimates of effect size: $After = 2.3$, ($95\% CI lower = 0.6, upper = 4.0$); $Before = -0.59$, ($95\% CI lower = -1.9, upper = 0.7$). *During* is the reference category.

Response Phase	Estimated Mean	Std. Error
Before	44.2	3.17
During	44.8	3.13
After	47.1	3.12

Table 4.1: *Estimated mean percentage Happiness during punchlines*

Credible performance: human vs. robot

A post-hoc GLMM linear model analysis of average percentage happiness of each audience member on each night with *Performer* (*Compere* / *HumanComedian* / *Robot*) as a fixed factor and *AudienceMember* and *Night* as a random factors shows a main effect of *Performer* ($F_{(2,227)} = 9.37, p < 0.01$). Planned pairwise comparisons show people responded more positively to the human comedian than the compère ($t_{227} = 4.33, p < 0.01$) but no other comparisons were significant (*Compere* vs. *Robot* $t_{227} = 1.85, p < 0.1$; *Comedian* vs. *Robot* $t_{227} = 1.37, p < 0.2$). As table 4.2 shows, people's responses to the robot were essentially intermediate between the two human performers.

Performer	Estimated Mean	Std. Error
Compère	38.2	4.1
Comedian	45.6	4.0
Robot	42.6	4.1

Table 4.2: *Estimated mean percentage Happiness per performer*

Gaze

A total of 14 individuals in the audience were randomly fixated a total of 153 times by the robot over the two nights. Three people were fixated only once and are excluded from the analysis.

The effect of *Gaze* on displayed *Happiness* is analysed in a GLMM linear model with *AudienceMember* as a random factor and *GazePhase* (2 seconds *Before* / *During* / 2 seconds *After*) a fixed factor. The robot's fixation points were not exact so the distance in pixels between a participant's actual location in the video and the fixation point of the robot is included as a co-variate. This analysis shows a main effect of *GazePhase* ($F_{(2,238)} = 14.5, p < 0.01$) and a main effect of *Distance* ($F_{(1,242)} = 5.19, p < 0.05$). Planned pairwise comparisons of *GazePhase* show no difference in the fixated person's displayed *Happiness* before and during the fixation (*Before* vs. *During* $t_{(238)} = 0.02, p = 0.99$) but a significant drop afterwards (*After* vs. *Before* $t_{(238)} = -4.68, p < 0.01$, *After* vs. *During* $t_{(238)} = -4.96, p < 0.01$). Fixed (*B*) coefficients for *GazePhase*: *Before* $= 5.8$, *During* $= 5.8$, *After* $= 3.2$. The fixed coefficient (*B*) for the distance co-variate of -0.12 additionally showed that the further an audience member's face was from the centre of the robot's fixation point the lower the estimated facial display of *Happiness*.

Gaze Phase	Estimated Mean	Std. Error
Before	45.1	4.3
During	45.1	4.3
After	42.1	4.3

Table 4.3: *Estimated mean percentage Happiness during robot gaze*

Pointing

A GLMM Linear Model analysis of average displayed *Happiness* in response to the point gesture with *ResponsePhase* (*Before* / *During* / *After*) as a fixed factor and *Night* and *AudienceMember* as random factors showed a significant main effect of *ResponsePhase* ($F_{(2,106)} = 6.11, p < 0.01$). The estimated means are provided in Table 4.4. Pairwise comparisons show that displayed *Happiness* increased during and immediately after the production of the gesture but were not reliably different while the gesture was produced and immediately after: *Before* vs. *During*: $t_{(106)} = -3.23, p < 0.01$, *Before* vs. *After* = ($t_{(106)} = -3.1, p < 0.01$), *During* vs. *After* ($t_{(106)} = 0.44, p = 0.66$). Fixed (*B*) coefficients: *Before* $= -8.7$, ($95\%CI lower = -14.4, upper = -3.14$); *During* $= 0.44$, ($95\%CI lower = -3.8, upper = 6.0$). *After* is redundant.

Response Phase	Estimated Mean	Std. Error
Before	42.3	4.3
During	52.2	4.1
After	51.1	4.0

Table 4.4: *Estimated mean percentage Happiness during point gesture*

4.3.6 Discussion

At the broadest level this study shows live performance experiments are viable for the study of live events. The intervention was unusual – the comic was plainly a robot, and a somewhat stilted one at that – but the results demonstrate that the performance was credible and provide the first objective evidence of individuated experience within an audience.

Displays of positive affect peaked just after the robot delivered punchlines. This shows people were selectively responsive to the content of the performance. By analysing the change in people's displayed emotions across the delivery of the punchline issues of possible bias in response to a robot are mitigated.[10] This demonstrates the credibility of the performance as a *comic* performance. Displays of positive affect during the robot's performance were essentially intermediate between the displays during the two performances by professionals. This further demonstrates the credibility of the performance, and the event more generally. In both these cases, the result is not quite the same as saying the audience laughed at the jokes, as our instrumentation captures a gradation of facial expression rather than laughter specifically.[11] Nonetheless, what is being measured here – effectively a smile – appears a good index of people's responses.

The audience responses were closely co-ordinated with the robot's gaze and pointing gesture: displays of positive affect declining after gaze, and increasing during and then holding after the gesture. This shows that people were selectively responsive to the content (punchlines, above) and the delivery (here). As the delivery manipulations are designed for individuals within the audience, this is important evidence of how an individual's experience of a live event will differ from others depending on the specifics of the social interactions they engage in, both within that event and across ostensibly identical showings.

An artefact of the robot system and the audience instrumentation allows further evidence for the effect of gaze. While the audience member's face is dynamically tracked, the robot can only be posed into preset positions. This means that to fixate on an audience member, the nearest gaze pose preset is chosen. To achieve credible gaze fixation a fixed seating plan was chosen, so that gaze poses appropriate for each seated position could be established prior to the show. The SHORE coordinates for the placeholder audience member used to find the gaze pose are known, and can then be compared to the coordinates for the actual audience member on the night. This essentially gives a measure of degree of eye-contact. The results show that the more closely the audience member was fixated by the robot, the more positive affect they displayed. This is consistent with the literature

[10] Discussion of audience bias given the use of a robot, and issues pertaining to embodied robotics more generally are reported in Katevas et al. (2015). Consideration of such aspects would be a diversion here given robots do not feature beyond this exploratory study.

[11] This would have been possible if the automatic results were correlated with manual annotations of laughter. However due to technical limitations at the time if the video stream was being analysed in real-time, as was necessary for the experimental manipulation, it could not also be archived, and so the source video is unavailable for analysis. Alternatively, further means of instrumentation could be found to capture the phenomenon directly.

of section 3.1.4, *Actions in the presence of others*, and is further evidence (to e.g., section 1.2.2, *Observing interaction*), that performers modulate audience responses not only through large performative gestures but also through the use of fine-grained mechanisms such as eye contact.

The responses to gaze and pointing are consistent with the hypothesis that the robot will be treated as a social agent and so will elicit response patterns typical of human interactions. Importantly, the results show this kind of interaction operating in the context of a live performance. This is consistent with the argument sustained through the previous chapters. At its simplest: in live events interaction matters. In particular, the results show that some of the interactional dynamics, well established for close, dyadic encounters extend to performers and individuals within an audience, despite the somewhat anonymised nature of massed audiences.

As an exploratory study, *Comedy Lab: human vs. robot* was a success. The robot work enabled experimental control and the instrumentation work enabled fine-grained measures of individuals' behaviour across the audience. There are some obvious limitations to the study: using a robot allowed for levels of control over aspects of delivery that would be unattainable with a human comedian but that obviously threaten the realism of the performance; this study provides data for only one audience and, as discussed above, each audience may have its own characteristics. In this case the stand-up comedians explicitly commented that the people who had gathered to watch a robot were not typical of a stand-up comedy audience. Nonetheless, this study provides clear experimental evidence, complimentary to that in section 2.1, *Observing stand-up comedy* of fine-grained social dynamics in operation that directly impact on people's experience of performance

The previous chapter asked how to identify the mechanisms and processes underpinning such social organisation at a sufficient level of detail. The instrumentation work in this chapter shows how this might be achieved. People's moment-by-moment reactions were explored by taking advantage of a technology that gathered these dynamic responses at a high temporal resolution. Note however, in this study the gaze data comes from only 14 people despite preparatory tests that indicated better coverage. This clearly could be improved, likely through more controlled lighting to ensure the quality of the source imagery. As suggested above, the instrumentation could also be better tailored to audience responses to stand-up comedy. Laughing, a distinct act, seems particularly amenable to instrumentation. Perhaps the accompanying head-bob or open mouth could be detected by further analysis of the computer vision data? Note also, that while the robot effectively self-reports gaze data, the gaze of the audience is unknown. Amongst other evidence throughout the previous chapter, mutual gaze was shown in section 3.2, *Performance?* to cue the listener's responses. This warrants instrumentation of audience gaze. While SHORE offers a "Rotation" measure of the face, but while this might have served as some kind of

approximation of gaze this has been found non-functional and other means are required.

Chapter 5

Experimenting with audiences, part one

The two main Comedy Lab experiments are presented in this chapter and the next. Having successfully gained evidence of a social effect of co-presence in the previous experiment, these two test the social effects of co-presence to the fullest extent practicable. This requires an expansion of the instrumentation, which is discussed first.

The experiment presented in this chapter manipulates performance. The basic premise is to have the performer an interacting party or not, and see what performer–audience and audience–audience dynamics are identifiable. The experiment contrasts live and recorded performance – directly addressing a topic that animates so much of the debate identified in chapter one.

The results provide good evidence for social dynamics within the audience, but little evidence for performer–audience interaction. These audiences appear indifferent to live vs. mediated performance, though high inter-subject variance renders this somewhat inconclusive given the between-subjects design. Nonetheless, these events were demonstrably social-spatial environments with heterogeneous audiences. The results emphasise that both conditions are live events, as regardless of the manipulation of the performance, the audience is live throughout.

5.1 Instrumenting an audience, redux

Instrumenting the audiences of *Comedy Lab: human vs. robot* through computer vision techniques worked. However as section 4.3.6 shows, the instrumentation should be expanded. Audience gaze should be captured. The instrumentation should be better tailored to audience responses to stand-up comedy. The coverage of the computer vision technique needs improving. This section addresses this, establishing the instrumentation for the experiments presented here and in the following chapter.

Gaze Analysing displays of happiness when the robot fixated its gaze on each audience member led to the most important result of the *Human vs. Robot* experiment. The robot effectively self-reports gaze data, but the gaze of the audience is unknown, as would be the gaze of a human performer. Capturing gaze data for all present thus presents a compelling analytic opportunity. In particular, mutual gaze could then be established between performer and audience member, (an important part of the interactional processes detailed in section 3.2, *Performance?*), and the lack of evidence for interaction within the audience be addressed (see section 3.3, *Audiences?*).

To track gaze in practice, the movement of the eye needs to be tracked relative to the head, and the head needs to be tracked relative to the room. Eye tracking equipment is well established, often based around cameras that track the position of the pupil. Variants that allow free movement of the head place the equipment on the head directly, however, such techniques are not suitable for this research, as cameras, electrodes, etc. surrounding the eyes could mask or disrupt any social display (see section 4.1, *Methodological issues*). Head tracking equipment is also well established, as a specific application of motion capture technologies. Markers of some form are placed on what is to be tracked, and the markers' positions are either triangulated using external equipment or they sense their own movement. Markerless technologies also exist although they tend to be less accurate. These analyse conventional video imagery or use depth sensing camera systems. For head pose at least, there are routes to instrument an audience.

Head pose is relevant irrespective of gaze. Head pose is mutually manifest to co-present people to a far greater degree than eye-gaze; experiments have shown that to judge eye gaze direction reliably people have to be looking straight at the person (up to an eccentricity of 4°) whereas judgments of others' head orientation are reliable even out to the extremes of peripheral vision (up to 90° eccentricity) (Loomis, Kelly, Pusch, Bailenson & Beall, 2008). Head pose has also been directly shown a significant cue in multi-party interaction; e.g., analysing a corpus of triadic interactions comprising video and upper body motion capture data, Battersby (2011) provides quantitative evidence that interlocutors' head and hand movements are organised by their dialogue

state and the task responsibilities that they bear. With less detail, Rienks, Poppe and Heylen (2010) shows a similar relevance of head pose in virtual reality recreations of meeting rooms. The salience of head pose should also apply within an audience, where the attention of others nearby can be judged through gross characteristics such as fidgety movement and heads clearly looking elsewhere than the stage – the kind of result found by Bailenson et al. (2008) in virtual reality recreations of classrooms where fellow students would e.g. look out the window or not, though the focus there was not on what cues were relevant (see section 3.4, *Motivating an experimental method*).

The Vicon MX motion capture system used by Battersby was available for this research, installed in the Media and Arts Technology Performance Lab at Queen Mary University of London. This is a marker-based system whereby the imagery from cameras around the room is processed to first determine the position of any marker in each two-dimensional image, and these positions along with calibration information of the cameras are then used to triangulate the position of each actual marker in three-dimensional space. In practice, this provides a capture volume of 8x4m on the floor raised up to standing height, within which any retroreflective coated 1cm ball will have its position tracked. The caveat is that this is subject to occlusion: there needs to be a line-of-sight between a ball and enough cameras to triangulate the position. In the context of a massed audience, with bodies packed tightly together and cameras placed high around the edge of the room, the only place with reliable sight lines is the top of the head. This is acceptable for the purposes here, allowing heads to be tracked if markers were placed on the top of the head. This can be achieved by fixing multiple markers onto hats that the people would wear. Multiple markers are required to obtain orientation of the hat (i.e. the relative positions of a rigid fixture of three or more points, not just the position of a single point).

The Performance Laboratory motion capture system provides viable instrumentation, given the constraint that the performance would have to take place within the capture volume in this particular room. Are there alternatives allowing instrumentation in existing performance venues that could scale to any audience size?

The instrumentation used for facial display of happiness could extend to provide head pose. This should be the most economic route in terms of equipment, and being markerless would not mask or disrupt social display. The rotation measure of SHORE™ shows an implementation of this, however it simply doesn't work. Coarser measures arrived at by manual annotation might be possible, as having many audience members in the one field of view suits at least one desired measure: *who is looking at who*.

Alternatively, autonomous devices that sense their own movement have the advantage that there is in principle no limit to the scale of the audience that could be captured,

and no infrastructure requirement of the venue. This incurs a fixed cost per unit, and re-introduces markers (or rather, a marker). These are typically small, solid state devices that contain an inertial measurement unit and storage to log the data to. As such, their sensing is relative rather than absolute, requiring 'homing' techniques to establish their absolute position and detect drift. Although now line-of-sight from any camera is not an issue, multiple units could be attached to instrument movement of other parts of the body; e.g., hands, given the consideration to gesture in the literature referred to in section 3.1, *Dynamics of interaction*.

Responses relevant to comedy Stand-up comedy audiences *laugh*. The existing measure of displayed happiness may capture laughter to some degree. However laughs manifest beyond facial display, most obviously in that we laugh out loud. Aural cues feature throughout the discussion of interaction in large audiences (section 3.3), and sound was used to motivate this research – being there at an event, in the moment: caught in the din of the crowd, or suddenly aware you could hear a pin drop (introduction, *Being there 'live'*). This presents a problem, as instrumenting audience individuals for sound has already been rejected as unfeasible.

Physiologically, laughter is a way of modifying breathing. Speaking is also, but laughter is more pronounced and breath control can be treated as its basic phonetic element (Kohler, 2008). Devices exist in medical practice to record breathing patterns by measuring chest expansion over time. While chest expansion is unlikely to be a social display itself, by instrumenting this core phenomenon it should act as a proxy for the wider ways laughter manifests, such as sound. Laughter across an audience can thus be instrumented by each person wearing these *breathing belts* around their chest and interpreting the data for pronounced exhalations. This instrumentation incurs a fixed cost per belt, plus the infrastructure to log the data. When worn, while it can be inconspicuous to others as it can be worn under outer clothing, it may make the wearer feel psychologically and/or physically uncomfortable.

A set of 16 Pneumotrace II™ breathing belts and an appropriate data logging device was available to this research, as used by researchers at University College London for e.g., Scott, Lavan, Chen and McGettigan (2014). Are there alternatives allowing instrumentation in existing performance venues that could scale to any audience size? Again, as laughter has visually manifest aspects, the instrumentation used for facial display of happiness could extend to index laughter. The chest may even be visible given the sight-lines of a camera position high in front of an audience, and breathing identifiable from the rise and fall of the relatively large area.

Reliability of computer vision analytics In the *Human vs. Robot* performances the coverage of audience members by computer vision analytics was less than anticipated.

Reviewing the documentation, it seems the video imagery of the audience is too dark for the analytics to identify most faces in the scene (e.g. see Figure 4.5, *Computer vision analytics running live*). While lighting levels had been verified beforehand during set-up, they are clearly too low in the actual performances. There are many factors that could account for this, for instance ambient lighting of the public space around the audience area being out of the experimenters' control, the lighting operator for the performance adjusting what had been preset, or that preset being lost, and so on. For these experiments, more control is needed, over both the venue and lighting equipment.

The bobbing of a head, orientation to others, smiling, laughing: together computer vision analytics, motion capture and breathing belts should allow these behaviours within an audience to be gathered at a high temporal resolution. Each of these systems have been used successfully in prior research. The constraint is the Media and Arts Technology Performance Lab, being where the available motion capture system is. This facility can accommodate the audience size set by breathing belt availability. In this facility, lighting and the surrounding spaces area entirely under the experimenter's control. The Performance Lab and these systems are suitable for the expanded instrumentation required for the main Comedy Lab experiments.

5.2 Comedy Lab: live vs. recorded

This Comedy Lab experiment is designed to provide a simple, minimal contrast between live, in-the-flesh performance and media playback of the same performance. Do people respond in systematically different ways to live versus recorded performance? The argument so far suggests people should, as there can be no performer–audience interaction with a recording, and interaction with the performer has been shown to factor in at least some performances. However little is known of the performer–audience and audience–audience dynamics beyond the addressor–addressee relationship in the evidence gathered so far.

Comedy Lab: live vs. recorded performance is a comedy gig with a five-minute warm-up routine delivered by a compère and fifteen minute stand-up routine delivered by the performer, to an audience of 16 participants each, arranged in a 4x4 grid. The mediation of performance is manipulated and the responses of individuals within the audience are compared. The presentation and content of the performance are held constant, while the audiences experience either *Live* or *Recorded* delivery. Specific question are detailed below.

Intense experience The accounts of live events in section 1.1.4, *Liveness?*, had intensified experience as a commonality. An intensified experience of stand-up comedy could be expected to result in longer laughs, harder laughter, or perhaps shorter-but-sharper responses. Longer or harder laughs would suggest the person finds the performance

funnier live. Shorter, sharper responses would suggest the person is more closely synchronised with the live performer's actions. Do the audiences in the live condition exhibit such differences?

Sense of others The other main commonality identified in section 1.1.4, *Liveness?* is having a sense of others. Being aware of others in the audience orienting to you may then systematically affect your responses to the performance, as may orienting to others in the audience.

Social encouragement A 'good' audience member would display positive affect to the performer as social encouragement, as suggested by section 3.2, *Performance?* and as shown in the eye-gaze results of the robot study (see section 4.3.6). This could be manifest in an overall effect of visibility to the performer – i.e., distance between performer and individual – or in moments when the performer is oriented to the individual and/or when the individual is oriented to the performer.

5.2.1 Design

To test people's responses to live versus recorded performance, one audience body was recruited, which was then split into two audience groups for a between-subjects design. Each audience group experienced the same performance – subject to the experiment's manipulation of mediation – in the same venue, within the same hour, with the same briefing and both having the compère deliver a warm-up act prior to the performance itself.

5.2.2 Materials

A performer-audience staging following the conventions of stand-up comedy was created in the 9x5m 'black-box' theatre space theatre space of the Media and Arts Technology Performance Laboratory at Queen Mary University of London. A 1.2x1m stage was created for the performer, facing a fixed arrangement of 16 chairs for the audience. The chairs were spaced out in an even grid, with rows centred ~1.2m apart and seats ~0.9m apart. A spotlight lit the stage, and a *Comedy Lab* logotype was transferred onto the wall behind. The arrangement of seating and stage can be seen in Figure 5.1.

For the *Recorded* performance a back-projected, person-sized screen was suspended over the stage at the same position as the performer in the *Live* performance. A recording of the performance was played back on the 1x1.7m screen using a short-throw data projector and monitor speakers. This arrangement can be seen in Figure 5.1. The image of the performer was mapped to the same size and position as the actual person when performing. The turn-around time from the *Live* to *Recorded* stagings was under 10 minutes.

The visual effect in both conditions is similar: a brightly lit life-size figure in an otherwise dim auditorium.

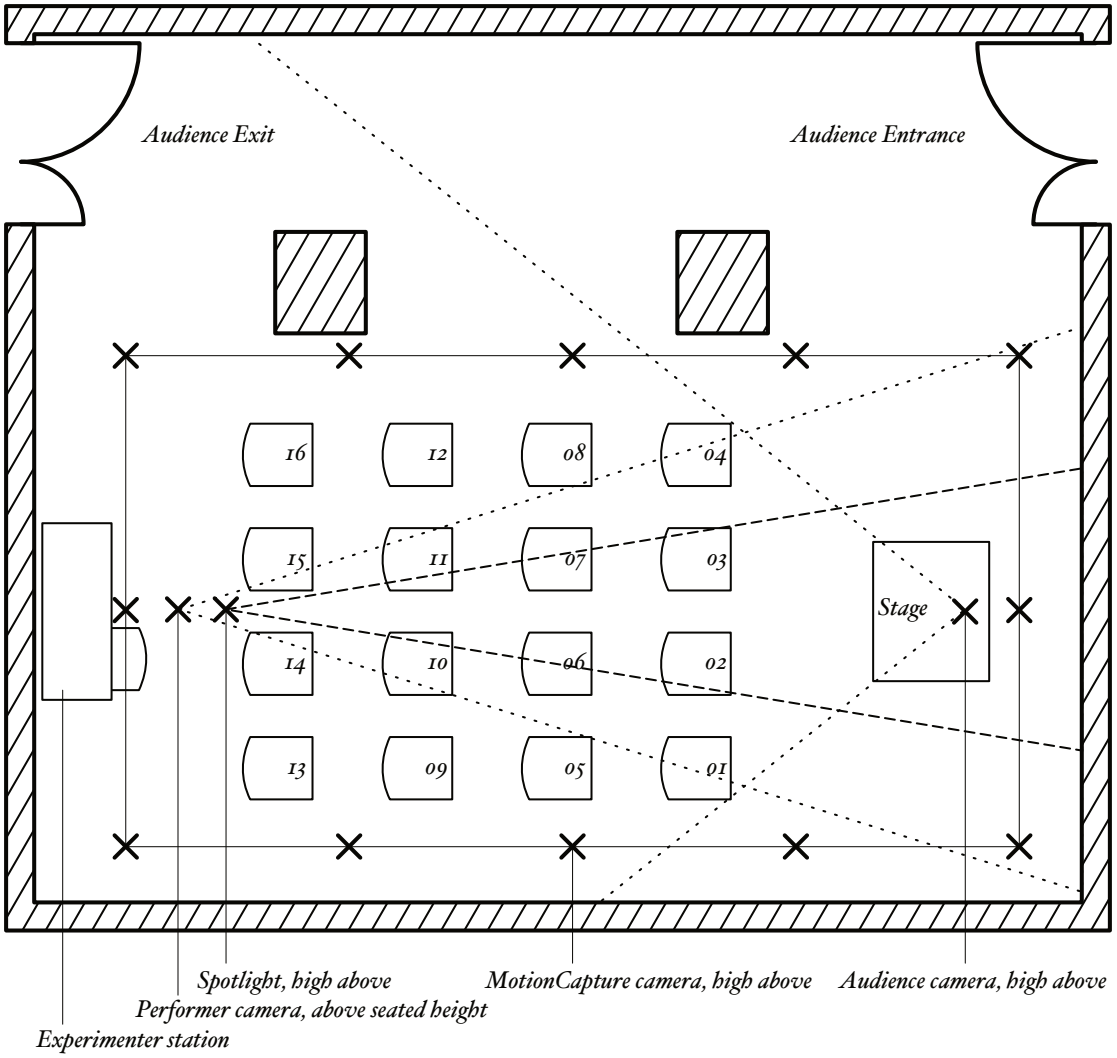


Figure 5.1: *Staging diagram for Comedy Lab: live vs. recorded. Live condition.*

5.2.3 Participants

Two professional stand-up comedians, Tiernan Douieb and Stuart Goldsmith, were recruited to carry out the performance, one as compère and one as the act. The compère was informed he would be performing a warm-up routine twice to two different audiences. The comedian performing the main act was informed his first performance would be recorded and played back to the second audience. Audience participants were recruited by advertising *Comedy Lab* through social media channels:

> *"Come and see some free stand-up comedy at Queen Mary, University of London. We are studying what makes a good performance. The comedian is booked, we need an audience of volunteers – you!".*

Correlating announcement and sign-up timings suggests the most effective channels were Queen Mary University of London's Facebook page and the compère and act's own Twitter accounts. 33 people signed-up online; 25 attended.

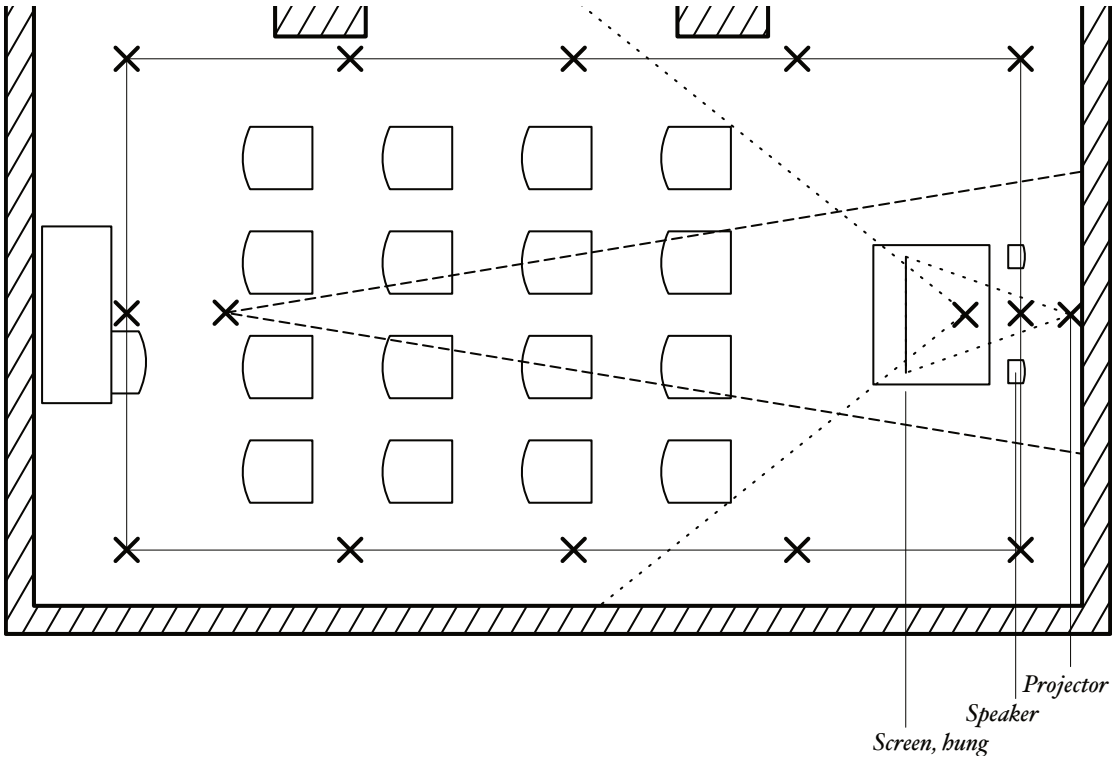


Figure 5.2: *Staging diagram for Comedy Lab: live vs. recorded. Recorded condition. Only differences from diagram 5.1 are annotated.*

Not having the full attendance of 32, for the live condition audience positions 04, 12, 15 and 16 were unfilled, and for the recorded condition audience positions 04, 13 and 16 were unfilled.

Participants were informed that they were being captured on video for research purposes and all data capturing and handling procedures were audited by the Queen Mary University of London research ethics committee (Reference: QMREC1199b).

5.2.4 Procedure

The audience assembled in an open room adjacent to the theatre space. A joint briefing was given by experimenter and compère. The briefing reiterated the recruitment message that what makes a good performance was being studied and stated that no personally identifiable materials will be published. The experimenter then formed two groups from the recruited audience. Sexes were kept balanced between the groups while otherwise the selection was random. The compère then led one group directly into the performance space where once the instrumentation was fitted the compère's warm-up started. On the main act finishing, the compère wrapped up the show, instructed the instrumentation to be removed, and led the audience group out via a different route to avoid mixing with the second audience group. The performance space was then turned-around for the recorded form of the main act, and once readied the compère led the second audience group in.

Figure 5.3: *Professional comic Stuart Goldsmith. Note hat with retroreflective motion capture marker fixture. This performer fixture is larger than audience fixtures (compare with Figure 5.4).*

5.2.5 Data capture

Each audience position was instrumented to individually record certain phenomena for the duration of the experiment:

Chest expansion Each participant was fitted with a wired device worn over their clothing around the chest that provides a measure of chest expansion. Chest expansion is thought to provide an independent, objective index of laughter responses (see section 5.1, *Instrumenting an audience, redux*). UFI Model 1132 Pneumotrace II breathing belts were used to capture this expansion, which feature piezo-electric respiration transducers that produce a linear signal in response to changes in thoracic circumference. The signals of these breathing belts were recorded by a single ADInstruments PowerLab 16/35 data acquisition system. Positions Audience 04 and Audience 16 failed to be captured.

Audio-Visual each audience position was within the shot of a high-definition (1080P25) video camera, positioned to have as similar a view of the audience as the performer while being hung from the lighting rig, so as not to distract from the performance. From the recorded video image subsequent analysis of the facial displays made by audience members from the point of view of the performer can be made by manual

Figure 5.4: *Comedy Lab: live vs. recorded. Photo of live condition, mid-performance. Note motion capture fixtures on audience members (see section 5.2.5,* Data capture*).*

annotation and computer vision techniques. Position Audience 01 failed to be captured.

Head pose each participant wore a cap modified such that its position and orientation could be tracked through the space. This was achieved by having the performer-audience staging within the capture volume of a Vicon MX motion capture system, comprising 12 near infra-red cameras hung from the lighting rig around the periphery of the space, and a fixture fitted with three retroreflective markers attached to the caps. Positions Audience 13-16 failed to be captured.

The performer was instrumented for head pose in the same way as the audience. An audio-visual recording was made to capture the performance from a generalised audience position. In practice the best position for the camera was behind the last row of audience and raised above their heads. Recorded audio was localised to the performer via a tie-clip radio-link microphone. The limited stage size constrained the performance within the capture volume of the motion capture system and field of view of the camera.

The data capture of the performer runs for the first, live, performance, and then forms the basis of the second, recorded, performance. Subsequent verification of the data capture shows performer head position remaining within the bounds of the screen (subject to

Figure 5.5: *Comedy Lab: live vs. recorded. Photo of recorded condition, mid-performance.*

inherent 3D to 2D assumptions). Some gestures extended beyond the bounds of the screen. The screen did not extend fully to the floor (which was out of the sight line of all but the front row), the image was cropped above the feet. There were five momentary pauses by our playback system during the recorded condition (1.7s-3.1s). This adds 10s to the 14m54s performance, and so to the duration of our (continuous) audience response data. Any measure that include performer data are absent in our data set during these pauses. Temporal alignment of audience responses to performance is otherwise verified through audio-sync between recordings.

5.2.6 Data set build

For each performance, the various data streams were compiled into a single dataset. A sampling rate of 10Hz was adopted as sufficient to capture the audience dynamics (e.g., the many mentions of 'tenths of a second' in the literature referred to in chapter 3, *Audiences and interaction*). This required the following tasks.

Audio-visual edits

The audio-visual recordings of performer and audience were edited together, placed side-by-side. Canonical dataset time is set by these edits.

The audience recording was found to fluctuate slightly in its timebase, and was retimed to ensure it was in sync with the performer recording; while their visual capture had some overlap, both recordings captured the audio of the performer sufficient for this to be used to achieve reliable synchrony. The recordings were filmed using identical cameras, however the audience recording was captured separately: motivated by giving the computer vision software uncompressed (i.e. full-fidelity) imagery, the audience recording was made through HDMI capture into a computer.

LaughState annotation

The audio-visual edits were then used to determine each audience member's *LaughState* through the performance. This is a manually annotated state of *Not* / *Smiling* / *Laughing* / *Indeterminate* and were coded by viewing the recording of the audience. The aim was to capture the simplest high-level categorisation of response state such as laughing or not. Reviewing the video, it became clear there was a third distinct state: smiling. Laughing was distinct in being dynamic in movement and extending beyond facial display to the body with i.e. bobbing of the head and movement of clothes around the chest. As these annotations are based upon the audio-visual edits, audience position 01 is missing.

Annotations were performed by the author in a custom application based upon *VCode* (Hagedorn, Hailpern & Karahalios, 2008) that presented an optimised user-interface for this task. While typically inter-analyst agreement is desired for manually coded data, here the data has not been verified by others as this indexing of the coarse phenomenon should be unproblematic given the inferential analyses ultimately performed.
Application: `http://tobyz.net/diary/2013/08/forked-video-annotation-app`
Source: `https://github.com/qmat/VCode-QMAT/`

Happiness – ShoreExporter

The video imagery recorded of the audience was used to generate measures of facial display for each audience member through the performance. This was broadly as per the exploratory study *Comedy Lab: human vs. robot*, described in section 4.2, *Instrumenting an audience*. The uncompressed audience video files were played through the SHORE demo application and a text file of raw measures generated for each. These raw data files were then processed in *ShoreExporter*. ShoreExporter then offset this timebase according to a single point of synchronisation, the start of the video. As these measures are based upon the audio-visual edits, audience position 01 is missing.

ChestExpansion – LabChartExporter

The ADInstruments data acquisition system produces a proprietary, binary data file. ADInstruments LabChart software was used to export this to a plain-text format suitable

for archiving and processing.

```
Interval= 0.001 s
ExcelDateTime= 4.1429633479525284e+004 04/06/2013 15:12:12.630984
TimeFormat= StartOfBlock
DateFormat=
ChannelTitle=   Channel 1   Channel 2   Channel 3   Channel 4   Channel 5   Channel 6   Channel 7   Channel 8   Channel 9   Channel 1
Range=          10.000      10.000      10.000      10.000      10.000      10.000      10.000      10.000      10.000      10.000
0               0.007       0.006       0.002       0.005       0.001       -0.014      0.001       0.000       0.004       -0.011
0.001           0.007       0.006       0.001       0.006       0.001       -0.014      0.000       0.000       0.005       -0.011
0.002           0.007       0.006       0.001       0.005       0.002       -0.014      0.000       0.000       0.005       -0.010
0.003           0.007       0.006       0.001       0.005       0.001       -0.014      0.001       0.000       0.005       -0.010
0.004           0.007       0.006       0.001       0.005       0.001       -0.015      0.001       0.000       0.006       -0.011
```

— *Excerpt of LabChart's plain-text output*

These raw data files were then processed in *LabChartExporter*, a tool patterned on ShoreExporter written by the author. The source data's timebase was corrected to dataset time according to a single point of synchronisation—a group breathe in and out—orchestrated by the compere in his warm-up.

LabChartExporter is a *Python 2.7* program of c. 150 lines. It performs the following tasks: i. transforms the timebase from 0.001s intervals to the 0.1s interval specified by the dataset by taking the arithmetic mean of the values recorded within each 0.1s period; ii. relabels the input channels to the corresponding audience member. The channel to audience position mapping etc. are set by a configuration file.

Usage: `python LabChartExporter.py input.log configuration.json`

LabExporter: `https://github.com/tobyspark/ComedyLab/blob/master/LabChartExporter/LabChartExporter.py`

Example configuration file: `https://github.com/tobyspark/ComedyLab/blob/master/LabChartExporter/configuration_labchart.json`

```
Time,   Audience01,   Audience02,   Audience03,   Audience04,   Audience05,   Audience06,   Audience07,   Audience08,   Audience09,
0.0,    0.007,        0.001,        0.004,        0,            0.006,        -0.014,       -0.011,       0.018,        0.002,
0.1,    0.00721,      0.00116,      0.00521,      0,            0.00553,      -0.01245,     -0.01051,     0.01733,      0.00102,
0.2,    0.00801,      0.00076,      0.00626,      0,            0.00506,      -0.01122,     -0.00987,     0.01703,      0.00126,
0.3,    0.00822,      0.00024,      0.00731,      0,            0.0058,       -0.01167,     -0.0061,      0.01606,      0.00148,
```

— *Excerpt of LabChartExporter output*

Moved, *OrientedAt* – **ViconExporter**

Motion capture data was recorded by the Vicon MX system's IQ software. During each performance, the IQ application would visualise the approximate 3D scene in realtime and record the raw camera feeds for more accurate post-processing free of any realtime constraint. In subsequent use, a 3D scene of markers could indeed be reconstructed from the camera data recorded for each performance, but the workflow enforced by IQ was then found incapable of transforming the individual markers into representations of the fixtures they were grouped upon. Three steps are necessary for this: i. individually identifying each virtual marker to its original; ii. creating virtual fixtures (rigid bodies) out of each group of

three markers that comprised a physical fixture; iii. computing the position information of each virtual fixture's three markers to give position and orientation (the full six degrees of freedom) of the fixture. Only Vicon's newer Blade software could support this given only the recordings. Once obtained, Blade was used to perform the post-processing tasks to transform the Vicon MX camera data into a file describing the six degrees of freedom information for each fixture throughout the performance.

Post-processing the raw Vicon data took considerable time. Markers 'drop out' of the reconstructed scene and so any identification (labelling) is lost when they reappear. While the software offers many automated techniques for the recovery of the markers' identity, it was found in practice that this often caused errors elsewhere that could take some time to notice and engineer a fix. The better use of time was just to proceed manually, scrubbing through time for each marker, drop-out by drop-out.

Vicon post-processing complete, the three performance scenes were exported as Vicon .V files. This is a proprietary, binary data file that required conversion to a plain-text format suitable for archiving. Processing requires an environment capable of manipulation of 3D geometries. A colleague had MATLAB scripts and expertise to read in .V files and analyse 3D data, and so this route was chosen.

ViconExporter is a collection of *MATLAB* functions of c. 500 lines to extract the data and c. 200 lines to calculate the measures. Producing the output from these functions and the parameters to correctly interpret the Comedy Lab recordings takes a further c. 300 lines. The data extraction aspect performs the following tasks: i. parses the binary '.v' file; ii. transforms the timebase from 50Hz to the 10Hz specified by the dataset by subsampling; iii. applies an offset to align the fixture's pose to a desired pose measure of the participant (ie. head straight forward); iv. applies dataset time; iv. writes comma-separated value text file of its output.

Creation of the offsets required a bespoke tool. The video recordings were augmented with the captured 3D information in an interactive application where the displayed pose direction could be 'nudged' into correct alignment. The application would then generate MATLAB code to define these rotations in the idiom of ViconExporter (see section 7.2, *Comedy Lab Visualiser*).

The measure calculations perform a geometric test between all participants to determine who is *oriented at* who, exports that data for visualisation, and uses that data to produce measures for statistical analysis such as *Moved* and *IsOrientedAtAudience*. Christopher Frauenberger wrote the initial implementation, which has been developed by the author since.

ViconExporter: `https://github.com/tobyspark/ComedyLab/tree/master/Vicon%20Exporter`

Usage: `https://github.com/tobyspark/ComedyLab/blob/master/Vicon%20Exporter/Vicon%20Exporter%20Readme.md`

Dataset info: https://github.com/tobyspark/ComedyLab/blob/master/Data%20-%20Raw/Motion%20Capture/Comedy%20Lab%20Mocap%20Dataset%20Info.txt

Time,	Audience01Hat/x,	Audience01Hat/y,	Audience01Hat/z,	Audience01Hat/rx,	Audience01Hat/ry,	Audience01Hat/rz,	Audience01H
249.7,	2369.5,	1418.7,	1163.4,	0.00010686,	-0.20381,	0.80977,	0.83502,
249.8,	2369.9,	1418.8,	1163.4,	0.0011235,	-0.20292,	0.80833,	0.83341,
249.9,	2370.1,	1418.9,	1163.3,	-6.9008e-06,	-0.20442,	0.81153,	0.83688,
250,	2370.5,	1418.9,	1163.3,	-0.0019867,	-0.20472,	0.80964,	0.83512,

— *Excerpt of ViconExporter pose output*

Time,	Performer_Hat_at_Performer_Hat,	Performer_Hat_at_Audience_01_Hat,	Performer_Hat_at_Audience_02,	Performer_Hat_at_Audience_03_H
476,	0,	0,	1,	0,
476.1,	0,	0,	1,	0,
476.2,	0,	0,	1,	1,
476.3,	0,	0,	0,	0,

— *Excerpt of ViconExporter OrientedAt output*

Time,	Audience 01/Moved,	Audience 01/Rotated,	Audience 01/isLookingAtPerformer,	Audience 01/isLookingAtAudience,	Audience 01/isB
297.6,	0,	0,	1,	0,	0,
297.7,	0.74243,	0.001698,	1,	0,	0,
297.8,	0.60125,	0.00031559,	1,	0,	0,
297.9,	0.57876,	0.0017269,	1,	0,	0,

— *Excerpt of ViconExporter measures output*

StatsExporter

With the various data captures conformed by the processes and exporter tools discussed above, the data for each modality recorded are ready for composition into a final, multi-modal dataset. The design of this dataset is informed by the needs of statistical analysis.

StatsExporter is a Python 2.7 program of c. 430 lines. It produces a comma-separated values text file listing each desired measure for each participant at each timestep through a performance.

Usage: `python StatsExporter.py configuration.json`

StatsExporter: https://github.com/tobyspark/ComedyLab/blob/master/Stats%20Exporter/StatsExporter.py

Example configuration file: https://github.com/tobyspark/ComedyLab/blob/master/Stats%20Exporter/Perf%203%20Data/Performance%203%20StatsConfig.json

AudienceID,	TimeStamp,	Light State While,	Laugh State,	Breathing Belt,	Happy,	Sad,	Surprised,	Angry,	MouthOpen
Audience 01,	472.0,	Unlit While None Lit,	Indeterminate,	-0.00036,	n/a,	n/a,	n/a,	n/a,	n/a,
Audience 01,	472.1,	Unlit While None Lit,	Indeterminate,	-0.00071,	n/a,	n/a,	n/a,	n/a,	n/a,
Audience 01,	472.2,	Unlit While None Lit,	Indeterminate,	-0.00096,	n/a,	n/a,	n/a,	n/a,	n/a,
Audience 01,	472.3,	Unlit While None Lit,	Indeterminate,	-0.00057,	n/a,	n/a,	n/a,	n/a,	n/a,

— *Excerpt of StatsExporter output*

5.2.7 Results

A dataset was assembled for the *Live* and *Recorded* performances, with the following measures for each audience member at 0.1 sec intervals during the performance. Each performance is deemed to have started just after the performer has walked on to stage, and to have stopped just before walking off the stage.

Laugh state A manually annotated state of *Not* / *Smiling* / *Laughing* / *Indeterminate*.

Happiness display A computed measure by the SHORE software performing sentiment analysis on the video recordings.

Chest expansion The voltage produced by the breathing belt. Positive values show expansion, negative show contraction. The timestamped value used for analysis is the mean of the 100 samples taken since the last timestamped value, 0.1s prior.

Movement The distance of the head pose point's translation, in mm, during one dataset frame (0.1s).

Oriented to performer *Yes* / *No*, computed from the head pose data. Inter-subject orientation is tested by whether any targets' head position is within a 0.7m radius cylinder extended forwards from the subject's head position/orientation. This geometric test was arrived at by experimenting with visualisation software developed specifically for this experiment (discussed further in chapter 7, *Visualising performer–audience dynamics*). The criteria was to judge whether the subject might appear, to any participant present, to be looking at the target. Note by this rationale a subject may be oriented at performer and one or more audience members simultaneously.

Oriented to screen *Yes* / *No*. An additional measure for the recorded condition. *OrientedToPerformer* for the recorded condition assumes a correspondence between subsequent (2D) screen representation and original (3D) capture. This measure does not, testing only what is materially present in the scene: the screen. Is the screen intersected by a line extended forwards from the subject's head position/orientation? Note such an empirically strict view precludes use of *OrientedByPerformer*.

Oriented to any audience member *Yes* / *No*, computed from the head pose data as above.

Oriented by performer Is the performer oriented to this audience member? *Yes* / *No*, computed from the head pose data as above.

Oriented by any audience member Are any audience members oriented to this audience member? *Yes* / *No*, computed from the head pose data as above.

Distance from performer A computed distance from performer's staged position to the centre point of the candidate seat, in meters.

Angle from performer A computed angle from performer's staged position to the centre point of the candidate seat, in degrees.

The datasets produced by the analysis have a resolution in tenths of a second and are therefore large with over 224,000 data points. This entails a trade-off in the analysis between the computational tractability of the analysing the entire dataset and the potential loss of information that occurs if the dataset is reduced by averaging over larger intervals. A two-stage approach is adopted to mitigate against this. The analysis begins with simple linear regression models that can be applied to the whole dataset. This is then followed by more focussed inferential analyses based on averages.

Patrick GT Healey helped with these statistical analyses.

Regression Analysis

In order to provide a simple comparison of measured *HappinessDisplay* in the live and recorded performances a linear regression was performed using the Automatic Linear Modelling procedure in SPSS with the all possible subsets method and automatic removal of outliers using Cook's distance. *Participants* (nominal), a binary encoding of *Mode* (*Live* vs. *Recorded*) and two continuous measures of position: *DistanceFromPerformer* (or screen) and *AngleFromPerformer* were entered as predictors. The resulting model produces an adjusted R^2 of 0.282 with only *Participant* as a significant predictor of displayed happiness ($F_{(1,16)} = 2,805$, $p < 0.00$, *Importance* $= 1.0$). The variation is marked, estimated means for individual displayed happiness vary between 11% and 81% over the course of the performances.

A parallel regression analysis for *Movement* produces a similar although much weaker result with an adjusted R^2 of 0.037. Only *Participant* is a significant predictor of movement ($F_{(1,16)} = 599.0$, $p < 0.00$, *Importance* $= 1.0$) with estimated marginal means varying between 1.01 and 4.10mm (in each tenth of a second window).

These tests provide relatively crude overall estimates of the effect of *Mode* on the dependent variables. In order to improve the model fit two separate regression analyses are performed that allow the inclusion of more fine-grained information about where participants were looking and who was looking at them. These variables are different for the *Live* and *Recorded* conditions because there is no embodied gaze-engaging performer in the recorded condition.

A linear regression was performed using the Automatic Linear Modelling procedure in SPSS with the all possible subsets method and automatic removal of outliers using Cook's distance. As before *Participants* (nominal) and two continuous measures of position:

DistanceFromPerformer and *AngleFromPerformer* were included as predictors. In addition for the *Live* condition the four binary orientation measures were included: *OrientedToPerformer*, *OrientedToAnyAudienceMember*, *OrientedByPerformer*, *OrientedByAnyAudienceMember*.

The resulting model for displayed happiness produces a slightly improved adjusted R^2 of 0.293. As before *Participant* is by far the most important predictor but there are also small but reliable effects of the orientation factors. The results are summarised in Table 5.1.

Source	F	df	Significance	Importance
Participant	1,603	9	< 0.001	0.976
OrientedByAudience	206.9	1	< 0.001	0.014
OrientedToPerformer	74.7	1	< 0.001	0.005
OrientedToAudience	66.7	1	< 0.001	0.005
OrientedByPerformer	3.04	1	< 0.081	0.000

Table 5.1: *Live performance: happiness*

The estimated marginal means show that being oriented to by another audience member correlates with significantly higher displays of happiness (estimated means: 41.6% rising to 48.5%) as does orienting to an audience member (estimated means: 42.7% rising to 47.4%). In contrast to this orienting to the performer correlates with a reduction in displayed happiness (estimated means: 46.5% falling to 43.6%) and being oriented to by the performer has no significant effect.

The parallel analysis for displayed happiness in the recorded performance substitutes orientation to the screen for orientation to the performer and has no factor of orientation by the performer. This produced an improved model fit with adjusted R^2 of 0.274. The results are summarised in Table 5.2.

Source	F	df	Significance	Importance
Participant	1,948.1	8	< 0.001	0.963
OrientedToScreen	271.8	1	< 0.001	0.017
AngleFromScreen	157.8	1	< 0.001	0.010
OrientedByAudience	122.4	1	< 0.001	0.08
OrientedToAudience	54.4	1	< 0.001	0.003

Table 5.2: *Recorded performance: happiness*

The basic response pattern is similar to that for the live performance. First, being oriented to by an audience member correlates with a reliable increase in displayed happiness (estimated means: 42.5% rising to 48.8%) as does orienting to an audience member (estimated means: 43.9% rising to 47.5%). Second orienting to the screen i.e., the performer's image, also correlates with a reduction in displayed happiness (estimated means: 48.2% falling to 43.3%) although the effect appears slightly stronger in the recorded case. The main difference is that the model for the recorded audience also shows an effect of

angle from the screen (performer) which shows an increase in displayed happiness with an increase in angle from the performer (Fixed coefficient = 55.2). This is mysterious.

The parallel regression analyses for the movement measure shows a poorer overall model fit. For the *Live* condition adjusted R^2 of 0.06. As before the largest effect is participant but there are also main effects of all the orientation factors.

Source	F	df	Significance	Importance
Participant	183.5	7	< 0.001	0.418
OrientedToPerformer	1,101.7	1	< 0.001	0.358
OrientedToAudience	345.4	1	< 0.001	0.112
DistancefromPerformer	182.1	1	< 0.001	0.059
OrientedByAudience	159.7	1	< 0.001	0.052
OrientedByPerformer	3.12	1	< 0.077	0.001

Table 5.3: *Live performance: movement*

People move less when they are orienting to the performer than when they are not; estimated marginal means 3.37 vs 2.07mm (in each tenth of a second window). They also move significantly less when they are orienting to another audience member than when they are not; estimated marginal means 3.79 vs 1.65mm. In contrast to this they move more when another audience member is orienting to them; estimated marginal means 2.21 vs 3.24mm. Note however that being oriented to by the performer does not have a reliable effect on movement. The effect of distance is that movement increases with distance from the performer.

For the recorded performance the overall model fit is again weak (R^2 = 0.04). As Table 5.4 shows there were main effects of all factors except being oriented to by other audience members.

Source	F	df	Significance	Importance
Participant	390.7	7	< 0.001	0.686
OrientedtoAudience	584.4	1	< 0.001	0.147
OrientedtoScreen	379.9	1	< 0.001	0.095
DistancefromScreen	193.3	1	< 0.001	0.049
AnglefromScreen	91.0	1	< 0.001	0.023

Table 5.4: *Recorded performance: movement*

As for the live performance people move less when orienting to the screen (performer's image) than when oriented elsewhere (estimated marginal means: 1.71 vs. 2.30mm) and less when oriented to another audience member than elsewhere (estimated marginal means: 1.07 vs. 2.94mm). The pattern for distance is also the same with movement increasing with both angle from the screen and distance from the screen.

Focused Analyses

The automated linear regression analyses can deal with large datasets but do not allow the construction of a model that specifies interactions between the fixed and random effects. To address this several Generalised Linear Mixed Model (GLMM) analyses were performed on aggregated datasets and fitting distributions that were appropriate to the data.

Laughter Coding In addition to the automatically collected data the videos were hand coded for whether speakers were overtly laughing, smiling or neither ('indeterminate' is ignored here).

As Figure 5.6 shows, the distribution of these hand coded states was different in the live and recorded conditions. A GLMM Gamma model analysis of time spent laughing with *Participants* as a random factor and *LaughState* (*Not* / *Smiling* / *Laughing*) and *Mode* (*Live* vs. *Recorded*) as fixed factors and *LaughState* $\times$ *Mode* as a two-way interaction shows no simple main effect of *Mode* ($F_{(1,61)} = 2.91, p = 0.09$) but a main effect of *LaughState* ($F_{(2,61)} = 37.7, p < 0.00$) and a *LaughState* $\times$ *Mode* interaction ($F_{(2,61)} = 37.7, p < 0.00$). This interaction, illustrated in Figure 5.6 arises because people laugh more in the *Recorded* condition and, correspondingly, do less not laughing. On average people laugh twice as much in the recorded performance as the live performance; Estimated marginal means (10ths of a second) Live = 722, Recorded = 1,942.

Another measure that was automatically gathered can be validated against the hand coding of laughter. Figure 5.7 suggests that the breathing belts appear to pick up patterns of overt laughter in the audience as marked by relatively large exhalations however, a GLMM Linear model analysis with *Participants* as a random factor and *LaughState* (*Not* / *Smiling* / *Laughing*) and *Mode* (*Live* vs. *Recorded*) as fixed factors and *LaughState* $\times$ *Mode* as a two-way interaction shows no simple main effect of *Mode* ($F_{(1,61)} = 0.07, p = 0.79$) nor a main effect of *LaughState* ($F_{(2,61)} = 2.17, p = 0.123$) and no interaction ($F_{(2,61)} = 0.66, p = 0.52$).

Orientation Following on from the regression analyses focussed GLMM analyses were performed with the average happiness and movement scores for each participant in each of the possible orientation states coded according to whether they were oriented to the performer (or screen) or audience and according to who was orienting them; another audience member or the performer. However, in the case of the recorded performance the performer's head orientation is notional – it refers to the counterfactual head angle calculated from the motion capture data for the live performance so as above we exclude oriented to by performer from the analysis that includes *Mode* as a fixed factor.

A GLMM linear analysis of average *HappinessDisplay* with *Participants* as a random factor, *Mode* (*Live* vs. *Recored*) and three binary orientation factors (*OrientedToPerformer*, *OrientedToAudience*, *OrientedByAudience*) as fixed factors and all their two and three-way

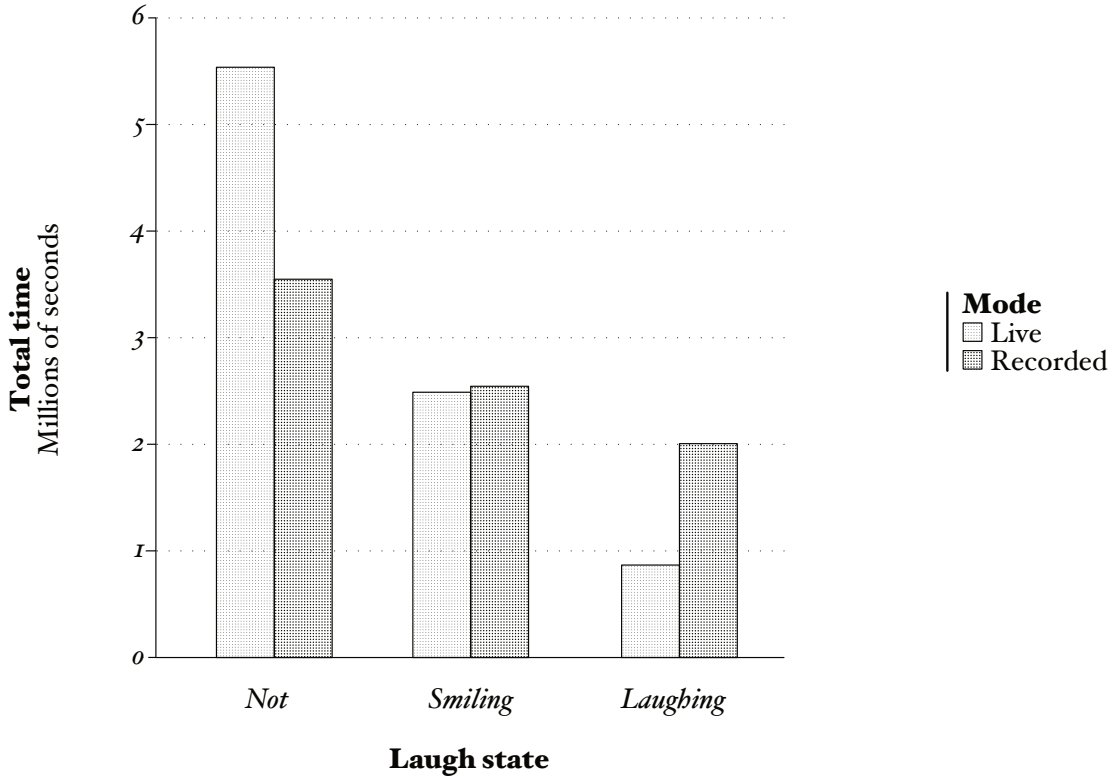


Figure 5.6: *Time spent laughing, smiling or not in live and recorded performances*

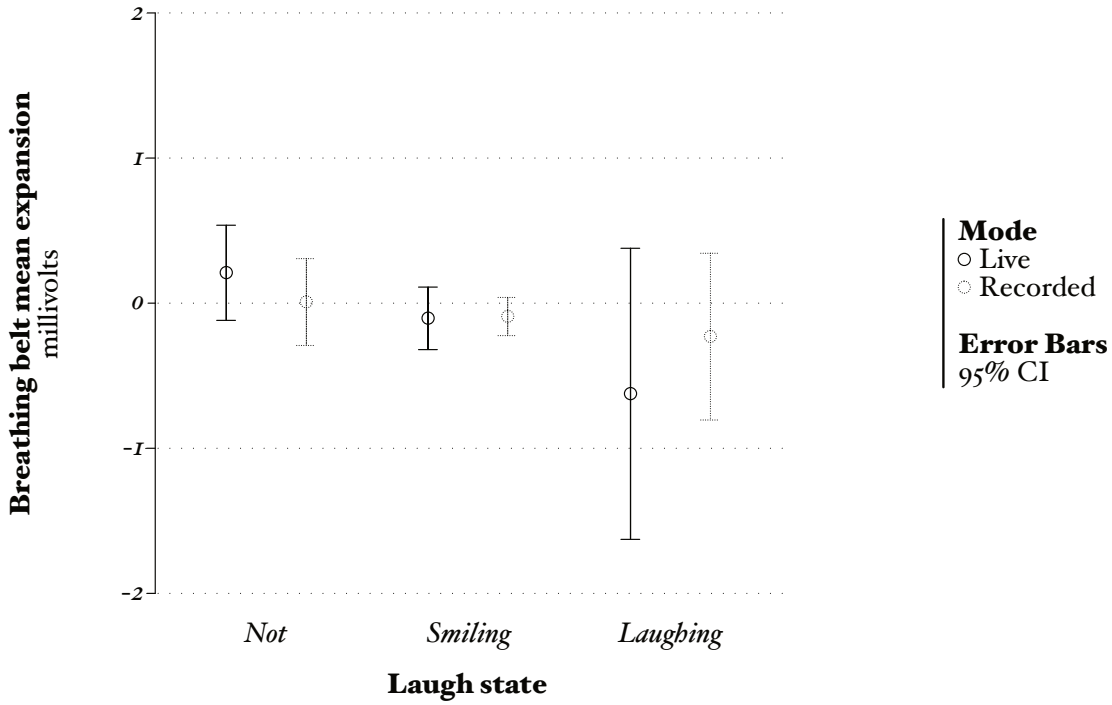


Figure 5.7: *Chest expansion when laughing, smiling or not in live and recorded performances*

interactions shows a main effect of orientation *OrientedToPerformer* ($F_{(1,113)} = 8.68$, $p < 0.00$) and *OrientedToAudience* ($F_{(1,113)} = 7.37$, $p = 0.01$). There is no reliable effect of either *Mode* ($F_{(1,113)} = 0.45$, $p = 0.51$) or *OrientedByAudience* ($F_{(1,113)} = 1.57$, $p = 0.21$). There were also no reliable interactions. This confirms the basic pattern from the regression analyses. In both the live and recorded performances people display higher levels of happiness when another audience member is orienting to them – regardless of where they are looking – and lower levels of happiness when they are orienting to the performer.

Orientation	Yes %	No %
To Performer	50.8	57.1
To Audience	57.2	50.7

Table 5.5: *Estimated marginal means for displayed happiness*

The parallel analysis for *Movement* using GLMM with a gamma distribution as the movement data are positively skewed shows a main effect of *OrientedAtPerformer* ($F_{(1,137)} = 105.6$, $p < 0.00$) but no simple main effects of *Mode* ($F_{(1,137)} = 1.66$, $p = 0.20$) nor *OrientedToAudience* ($F_{(1,137)} = 2.61$, $p = 0.12$). There is however a significant two-way interaction between *OrientedAtPerformer* and *OrientedToAudience* ($F_{(1,137)} = 15.4$, $p < 0.00$). As with the regression analysis, overall people move less when orienting to the performer (estimated marginal means 4.62 vs. 1.54). Note this correlation is unlikely to be due to the smaller target size of the performer as the measure is of translation rather than rotation; a person re-orienting within the differing bounds would 'bob' similarly (translation) but turn very distinctly. The interaction shows that when they are not orienting to the performer there is an effect of orientation to the audience. This follows the same general pattern as the regression analyses with people moving less when orienting to the audience than, in this case to neither the audience nor the performer. Movement here appears to be picking up on people shifting between orientations and is therefore highest when they are attending to neither the performer/screen nor each other.

Orientation	Audience Yes	Audience No
Performer Yes	1.72%	1.37%
Performer No	3.40%	6.27%

Table 5.6: *Estimated marginal means for movement interaction between orientation to performer and to any audience member*

5.2.8 Discussion

In the initial discussion of liveness, much of the theory concerned mediated performance. Notably, irreconcilable positions on the (in-)difference of live and mediated performance feature in section 1.2.3, *No interaction?*, and the subsequent discussion of event cinema in section 1.2.4, *No co-presence?*, can be read as a quasi-experiment to test the impact of live vs. mediated performance. This study carried out a controlled, experimental comparison of audience responses to live vs. mediated performance.

The results show that people laugh longer in the recorded performance, and their displayed happiness is stronger in the live performance. This appears to fit the idea of live events being intense: introducing the experiment, it was posited that an intensified experience of stand-up comedy could result in shorter and sharper responses. Rather than a live performer resulting in bigger laughs, this would have in effect 'the same amount of funny' expressed with closer synchrony to the live performer's actions. Close synchrony would fit the evidence of fine temporal coordination seen throughout section 3.1, *Dynamics of interaction*, and show wider applicability of the moment-by-moment processes seen in section 3.2, *Performance?*

However, none of the regression analyses or fitted models suggest any reliable difference between live and recorded performances. While shorter-and-sharper responses might be averaged out (and so lost) by such analyses, other aspects of interaction should manifest in ways that are explicit in these analyses. For example, the anticipated effect of distance or angle from the performer, as per the question of social encouragement in the introduction to this experiment. This factor is not found to be significant, however. Another is that when orienting to the performer (or screen) people move less but also smile less, which is not what social encouragement would suggest. Overall, their displays of happiness do not seem sensitive to the potential interaction the live performer. Based on this analysis audiences are indifferent to live vs. mediated performance. Whether this is a fair determination given the high intra-subject variance is arguable, however. A between-subjects design relies upon the two population samples to be broadly equivalent, and with the small sample sizes here people may just act too individually for a comparison to be made. A larger sample size may alleviate this problem, or the obstacles of a within-subjects design for such a contrast be overcome.

What is clear in this data is the degree to which the events were social-spatial environments with heterogeneous audiences. This is a position taken up most notably by Reason (2004) but also seen across the first chapter's discussion of liveness. Across the regressions the most obvious point is that individual differences between participants are the most important predictor of both displayed happiness and movement. Some people move more than others and some smile (a lot) more than others. The individual variation in degree of response is large and means that the tests for the effects of other factors are less powerful.

Nonetheless, there is a also a clear effect of orientation. People smile more and move less when they are orienting to another audience member. People's displays of happiness are thus clearly sensitive to their interactions with other audience members. The contrast between happiness displayed to other audience members and that displayed to the performer/screen is clear, suggesting that the primary axis of social interaction is, in fact, between audience members.

The performer orientation result at first sight contrasts with the results of *Comedy Lab: human vs. robot*; a positive effect of closeness of eye-gaze compared with a negative effect of being oriented towards by head pose. However they are not directly comparable as in the robot study the effects of specific gaze fixation on specific individuals was examined whereas in this case what is tested are the effects of general orientation to, in effect, regions of the room. There is a paucity of studies that compare head orientation and eye gaze in multiparty interaction with which to contextualise these results. One proposal is that multiparty dialogue is orchestrated by cues which sit on a continuum, with eye-gaze as rather ineffectual at one end, moving through head orientation and ending with highly effective gesture. This is made by Battersby (2011, p.150) drawing upon his corpus of motion capture data and the limited results of (Jokinen, Nishida & Yamamoto, 2010). The Comedy Lab results suggest that for live events, it is not this simple. Notwithstanding the basic psychophysical fact that head orientation is mutually manifest to co-present people to a far greater degree than eye-gaze, when you are held by that eye-gaze it is highly effectual.

There is a further comparison to the robot study to be made. This study is not quite the first manipulation of performance mode, as to establish the credibility of the robot's performance, audience responses were gathered to the human performer and compared. The result there is similar to the result here: that audiences are broadly indifferent to the mode of performance.

Comedy Lab: live vs. recorded provides good evidence for social dynamics in live events. Combined with the lack of direct evidence for the impact of a live performer, this draws attention to what is meant by not-live. While the recorded condition is ostensibly not live, in both cases there is a live audience and it is the audience–audience dynamics that are notable in this event, not performer–audience dynamics.

Chapter 6

Experimenting with audiences, part two

Building on the findings of *Comedy Lab: live vs. recorded* this chapter focusses on an experimental manipulation of the audience. Anecdotal evidence from our observations of street performance, cabaret and stand-up comedy suggests that an important factor influencing audience behaviour is the degree to which people feel they are visually exposed to other audience members. For instance, *Cafe de Paris*[1] features a mix of cabaret tables, booths and standing balconies, the different configurations appearing to shape audience behaviours. Another London venue with an unusual configuration is *Shakespeare's Globe*,[2] extensively studied by Woods (2012, p.98) who remarks that "its equally-lit open-air nature, and organisation almost in-the-round, promote a particular audience co-presence that informs behaviour and response". This experiment attempts to investigate these dynamics directly and to test their effects on people's social responses to each other and to the performer.

The results provide strong evidence for social dynamics within the audience, and limited evidence for performer–audience dynamics. The manipulation is of lighting, being spotlit reduces people's responses, while everyone being lit increases people's responses: it is the effect of being picked out not being lit *per se* that matters. The results affirm that live events are social-spatial environments with heterogeneous audiences.

[1] "The home of cabaret and dining in London since 1924, the iconic ballroom of Cafe de Paris hosts weekly Friday and Saturday night performances from the crème of the worlds of circus, burlesque, music, comedy and magic." — http://www.cafedeparis.com

[2] A modern reconstruction of the 1599 century Globe Theatre — http://www.shakespearesglobe.com

6.1 Comedy Lab: lit vs. all lit

This Comedy Lab experiment manipulates the exposure of audience individuals within the performance environment. *Comedy Lab: lit vs. all lit audience* is a comedy gig with a five-minute warm-up routine delivered by a compère and fifteen minute stand-up routine delivered by the performer, to an audience of 16 participants each, arranged in a 4x4 grid. The lighting condition across the audience is manipulated and individual audience responses compared. The audience variously experience no lighting, a randomised pattern of individual coverage, and complete coverage. Specific questions are detailed below.

Availability If you are lit, your visual displays are more available to others. Does everyone being lit rather than in the dark have an effect on people's responses?

Exposure If you are spotlit while those around you are in the dark, your visual displays are exposed to the others, whereas the displays of the others remain less available. Does being exposed have an effect on people's responses?

6.1.1 Design

The lighting state across the audience positions was manipulated during the performance. There were three phases of lighting: all low, individuals randomly lit, all high. This resulting in individual audience *LightState* of *Lit* / *Unlit*. . Lighting state was thus a within-participants factor. Order effects were spread across the three cycles of lighting.

6.1.2 Materials

The materials of 5.2, *Comedy Lab: live vs. recorded*, were used and extended with a lighting rig, described here. A spotlight was arranged above and slightly forward of each audience position with beam set to light the position alone. The lighting state at low had these spotlights set at the minimum brightness for the camera to obtain an image with audience facial expression visible. The lighting state at high had the lights set to maximum brightness to make as overt a change, and state, as possible.

These spotlights were under computer control, and a program was written to cycle the lights. For the individually lit phase, it was ensured that each audience position to light was randomly chosen within the constraints that each position was lit once in each phase, and one position was lit in each audience row simultaneously. Note each position being lit once in each phase follows the experimental design, doing so in rows simultaneously is a concession to getting through all positions quickly enough to allow repeating each phase within one performance. The three phases were cycled through three times during the performance. With the set performance time of 15 minutes, each phase lasted 1m 40s. A short (less than one second) fade was applied on each light between these states, as otherwise

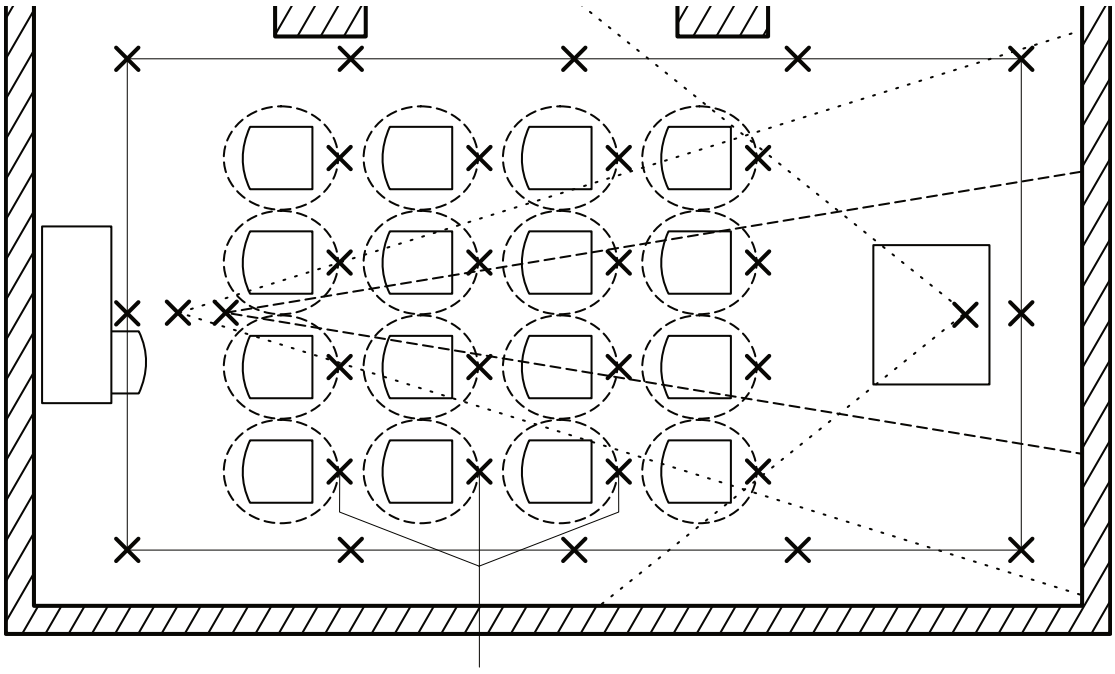


Figure 6.1: *Staging diagram for Comedy Lab: lit vs. all lit.*
Only differences from diagram 5.1 are annotated.

the change was instantaneous and seemed overly jarring to those present.[3] The lighting pattern recorded from the performance is shown is Figure 6.2 (here, the performer is to the left, so audience position 01 is top left of each 4x4 grid).

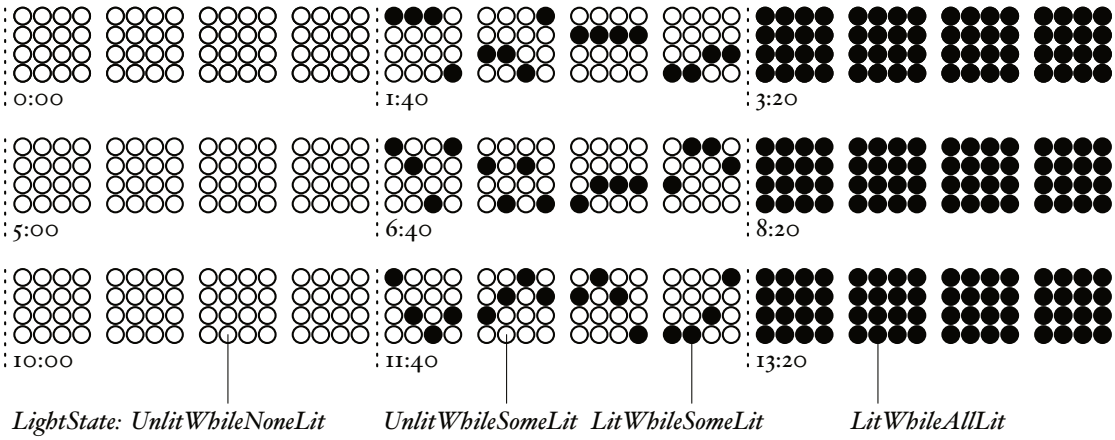


Figure 6.2: *LightState during the performance experiment.*

6.1.3 Participants

Recruitment was as reported for 5.2, *Comedy Lab: live vs. recorded*. 30 people signed-up online; 20 attended. This many people meant all instrumented audience positions were filled while

[3]Perhaps related: this made the behaviour of the new LED based lights more akin to that of traditional incandescent lights.

four stood behind.

Participants were informed that they were being captured on video for research purposes and all data capturing and handling procedures were audited by the Queen Mary University of London research ethics committee (Reference: QMREC1199b).

6.1.4 Procedure

As reported for 5.2, *Comedy Lab: live vs. recorded*, but with only the live performance. The lighting cycle detailed in the design section operated for the duration of this act.

Figure 6.3: *Comedy Lab: lit vs. all lit. Photo mid-performance with some audience members exposed in the LitWhileSomeLit condition.*

6.1.5 Data capture

As reported for 5.2, *Comedy Lab: live vs. recorded*, with the further step of capturing *LightState*. While the partially randomised lighting sequence was generated and run by a computer program, this program did not archive the lighting state through each experiment. Instead this state data was obtained from manual annotation of the audio-visual edits.

6.1.6 Dataset build

As per 5.2, *Comedy Lab: live vs. recorded*, with the further step of transforming *LightState* into *LitState*. This transforms the measure of each audience position's lighting state considered in isolation (ie. lit vs. unlit) to how it relates among all audience positions (ie. lit, as are all other positions): *LitWhileAllLit* / *LitWhileSomeLit* / *UnlitWhileSomeLit* / *UnlitWhileNoneLit*. This task is part of *StatsExporter*.

6.1.7 Results

A dataset was assembled for the performance, with the following measures for each audience member at 0.1 sec intervals during the performance. Further description for each measure is reported in the equivalent section (5.2.7, *Results*) of *Comedy Lab: live vs. recorded*. The performance is deemed to have started just after the performer has walked on to stage, and to have stopped just before walking off the stage.

Lit state A manually annotated and then computed state of *LitWhileAllLit* / *LitWhileSomeLit* / *UnlitWhileSomeLit* / *UnlitWhileNoneLit*.

Laugh state A manually annotated state of *Not* / *Smiling* / *Laughing* / *Indeterminate*.

Happiness display A computed measure by the SHORE software performing sentiment analysis on the video recordings.

Chest expansion The voltage produced by the breathing belt. Positive values show expansion, negative show contraction.

Movement The distance of the head pose point's translation, in mm, during one dataset frame (0.1s).

Oriented to performer *Yes* / *No*, computed from the head pose data.

Oriented to any audience member *Yes* / *No*, computed from the head pose data.

Oriented by performer *Yes* / *No*, computed from the head pose data.

Oriented by any audience member *Yes* / *No*, computed from the head pose data.

Distance from performer A computed distance from performer's staged position to the centre point of the candidate seat, in meters.

Angle from performer A computed angle from performer's staged position to the centre point of the candidate seat, in degrees.

As with *Comedy Lab: live vs. recorded* the datasets produced by the analysis have a resolution in tenths of a second and are therefore large with over 151,000 data points. This entails a trade-off in the analysis between the computational tractability of analysing the entire dataset and the potential loss of information that occurs if the dataset is reduced by averaging over larger intervals. Again, a two-stage approach is adopted to mitigate against this. The analysis begins with simple linear regression models that can be applied to the whole dataset. This is then followed by more focussed inferential analyses based on averages.

Patrick GT Healey helped with these statistical analyses.

Regression Analysis

Happiness display Using the Automatic Linear Modelling procedure in SPSS with the all possible subsets method and automatic removal of outliers using Cook's distance. *Participants* (nominal), two continuous measures of position: *DistanceFromPerformer* and *AngleToPerformer* and *LitState* (*UnlitWhileNoneLit*, *LitWhileAllLit*, *LitWhileSomeLit* and *UnlitWhileSomeLit*) were entered as predictors. The resulting model produces an adjusted R^2 of 0.44 with *Participant* as a significant and powerful predictor of displayed happiness ($F_{(1,12)} = 3,726, p < 0.00, Importance = 1.0$). The variation in individual responses is very marked with estimated means for individual *HappinessDisplay* varying between 5% and 75% over the course of the performances. In addition, there is a main effect of *LitState* ($F_{(1,3)} = 95.5, p < 0.00, Importance = 0.004$). Estimated means show that people display least happiness when individually spotlit (*LitWhileSomeLit* 31%) and most when everyone else is also lit (*LitWhileAllLit* 36%). Raw means are provided in Figure 6.4.

Movement *Participants* (nominal), two continuous measures of position: *DistanceFromPerformer* and *AngleToPerformer* and *LitState* (*UnlitWhileNoneLit*, *LitWhileAllLit*, *LitWhileSomeLit* and *UnlitWhileSomeLit*) were entered as predictors. The resulting model is a relatively poor fit with an adjusted R^2 of 0.025. *Participant* is a significant and powerful predictor of *Movement* ($F_{(1,7)} = 345, p < 0.00, Importance = 0.83$). Note: degrees of freedom reduced to 7 because participants whose movements do not contribute anything to the model are removed from it with this procedure. The variation in individual responses is substantial with estimated means for individuals varying between 1.0 and 2.7mm per frame over the course of the performance. In addition, there is a main effect of *LitState* ($F_{(1,3)} = 163, p < 0.00, Importance = 0.17$). Estimated means show that people move least when spotlit (*LitWhileSomeLit* 1.54mm per frame) and most when everyone is lit (*LitWhileAllLit* 2.2mm per frame). Raw means are provided in Figure 6.5.

Chest expansion *Participants* (nominal), two continuous measures of position: *DistanceFromPerformer* and *AngleToPerformer* and *LitState* (*UnlitWhileNoneLit*, *LitWhileAllLit*, *LitWhileSomeLit* and *UnlitWhileSomeLit*) were entered as predictors. The resulting model

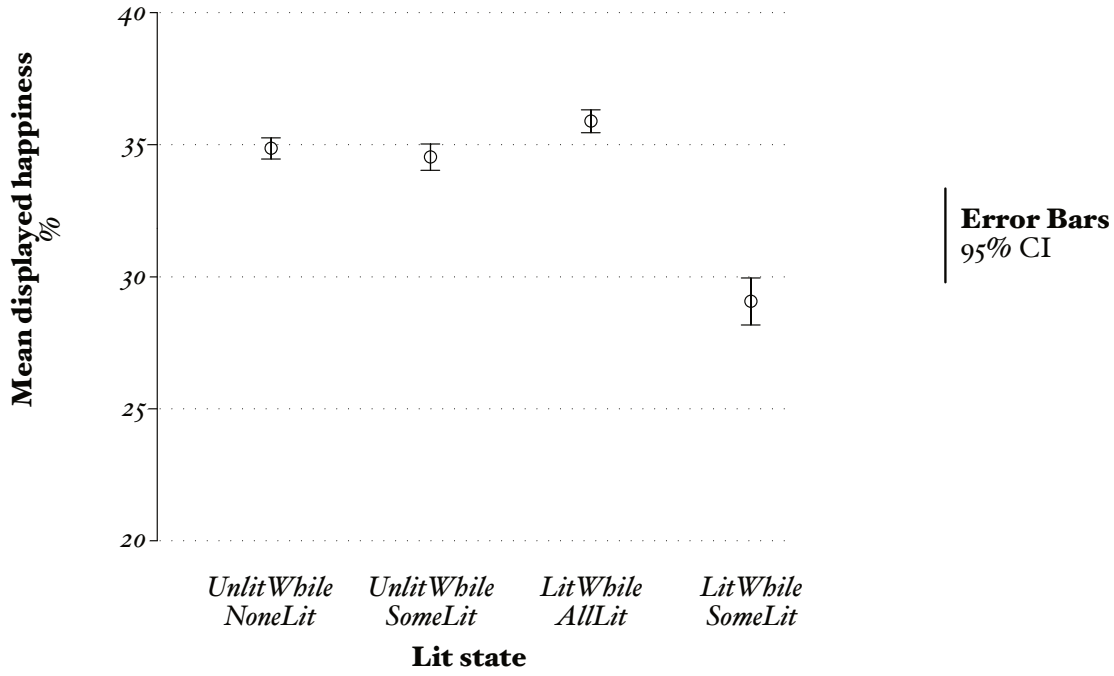


Figure 6.4: *Happiness display by lighting condition*

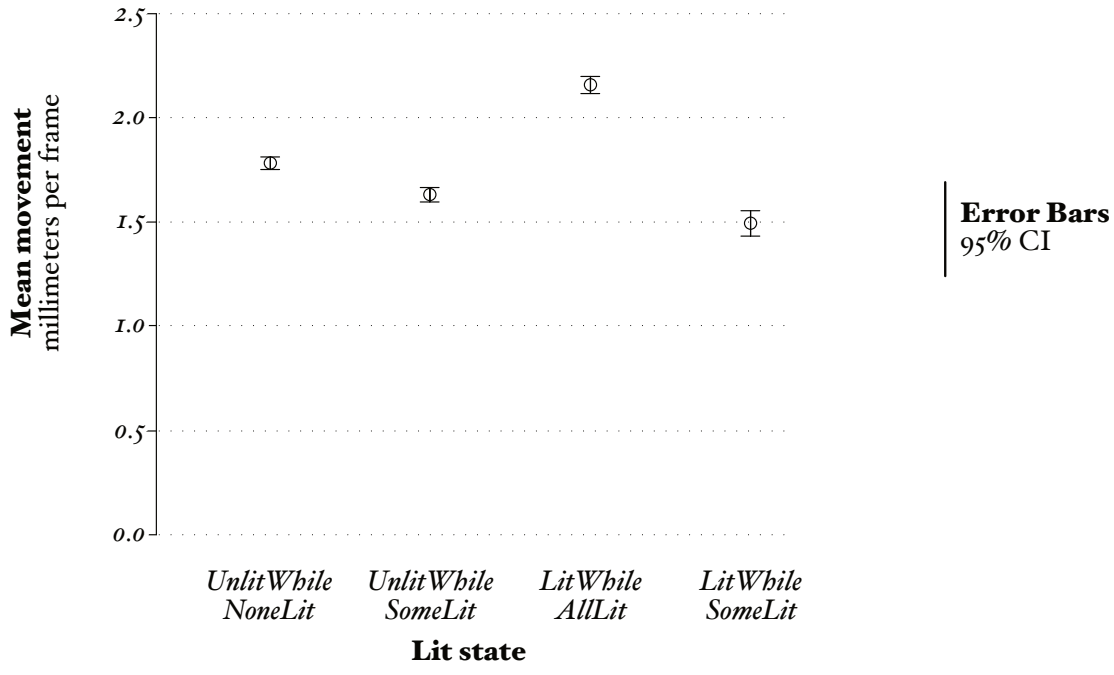


Figure 6.5: *Movement by lighting condition*

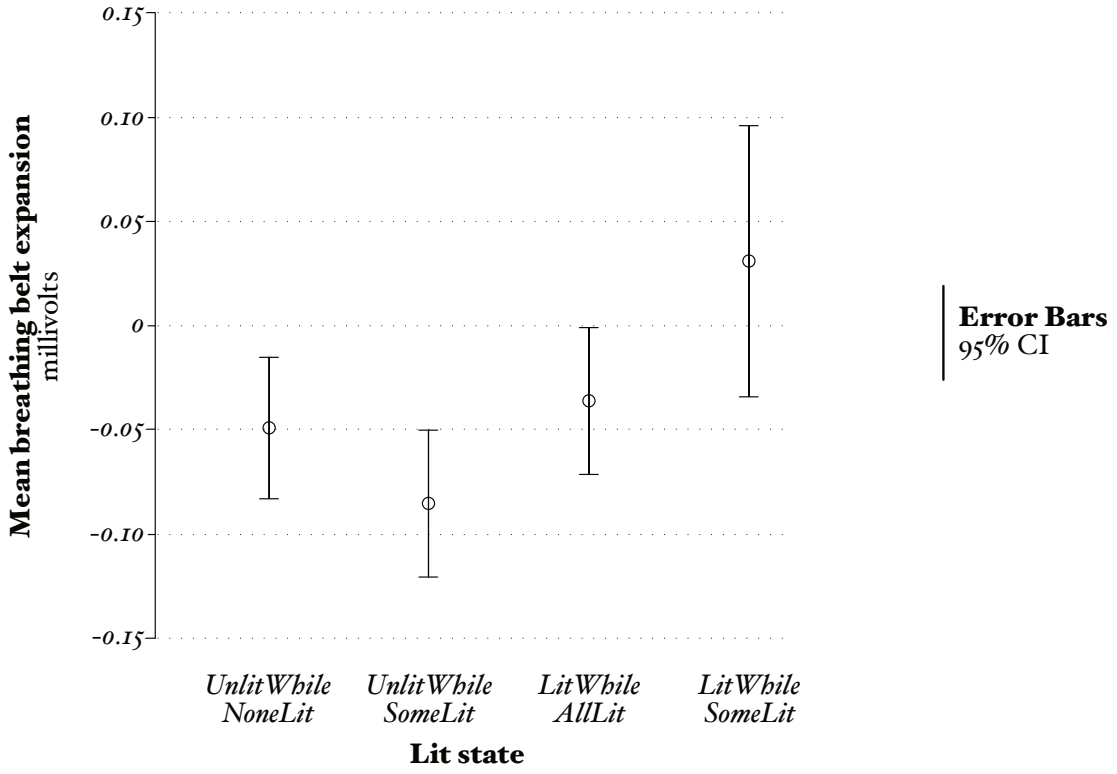


Figure 6.6: *Chest expansion by lighting condition*

is a relatively poor fit with an adjusted R^2 of 0.002. *Participant* is a significant and powerful predictor of *Chestexpansion* ($F_{(1,5)} = 64.8, p < 0.00, Importance = 0.97$). In addition, there is a main effect of *LitState* ($F_{(1,2)} = 4.38, p = 0.01, Importance = 0.026$). Estimated marginal means for individual breathing belt readings vary between 1920mV and -0.41mV. For the lighting comparisons estimated marginal means indicate the *LitWhileSomeLit* condition was -0.70mV and the *LitWhileAllLit* condition was -0.12mV . The raw means (note: not EMMS) are shown in Figure 6.6.

Focussed Comparisons

Happiness Display Average happiness for each participant in each lighting state was calculated and analysed in a GLMM linear analysis with *Participants* as a random factor and *LitState* as a fixed factor with four within-subjects levels (*LitWhileAllLit*, *UnlitWhileNoneLit*, *UnlitWhileSomeLit*, *LitWhileSomeLit*). This shows a marginal (but consistent with the other effects) overall effect of *LitState* ($F_{(3,56)} = 2.65, p = 0.06$). Planned pairwise comparisons of the four lighting states confirm the basic pattern illustrated in 6.4: people display less happiness when spotlit than when in the dark (*LitWhileSomeLit-UnlitWhileNoneLit* $t_{(56)} = -2.75, p = 0.058$) or than when in the dark but some others are spotlit (*LitWhileSomeLit-UnlitWhileSomeLit* $t_{(56)} = -2.04, p = 0.05$). The spotlit vs. all lit comparison is in the same direction but not significant (*LitWhileSomeLit-LitWhileAllLitt*$_{(56)} = -1.89, p = 0.07$).

Movement These data are positively skewed so a GLMM analysis with a Gamma distribution is used with *Participants* as a random factor and *LitState* as a fixed factor with four within-subjects levels (*LitWhileAllLit*, *UnlitWhileNoneLit*, *UnlitWhileSomeLit*, *LitWhileSomeLit*). This shows a overall effect of *LitState* ($F_{(3,44)} = 21.6, p < 0.00$). Planned pairwise comparisons of the four lighting states show that people tend to move most when all are lit and least when spotlit (*LitWhileAllLit-UnlitWhileNoneLit* $t_{(44)} = 2.00, p = 0.05$; *LitWhileAllLit-UnlitWhileSomeLit* $t_{(44)} = 4.25, p < 0.00$; *LitWhileAllLit-LitWhileSomeLit* $t_{(44)} = 6.49, p < 0.00$; *UnlitWhileNoneLit-LitWhileSomeLit* $t_{(44)} = 2.33, p = 0.02$. No other comparisons were significant). This confirms the basic pattern suggested by the raw means in Figure 6.5.

Chest expansion GLMM linear model analysis of the breathing belt data with *Participants* as a random factor and *LitState* as a fixed factor with four within-subjects levels (*LitWhileAllLit*, *UnlitWhileNoneLit*, *UnlitWhileSomeLit*, *LitWhileSomeLit*) shows no main effect of *LitState* ($F_{(3,60)} = 0.37, p = 0.77$) and no planned pairwise comparisons show significant effects. This is likely to be a consequence of the high variance in this measure. Difficult to conclude much.

Oriented to To check if any of these effects are influenced by people's orientation a second set of focussed analyses is performed. Average happiness and movement for each participant in each lighting state $\times$ each orientation state were calculated. Breathing belt data are dropped because preceding analyses suggest that the high variance makes it less suitable as a measure. To simplify the analyses each orientation factor is considered separately.

A GLMM analysis of *HappinessDisplay* using a gamma distribution (which in this case gives a better fit than a linear distribution: Akaike Corrected Information Criterion = 360.5) with *Participants* as a random factor and *LitState* and *OrientedToPerformer* (Y/N) as fixed within-subjects factors shows a main effect of *LitState* ($F_{(3,201)} = 26.2, p < 0.00$) no simple main effect of *OrientedToPerformer* ($F_{(1,201)} = 1.01, p = 0.32$) but also a reliable *LitState* $\times$ *OrientedToPerformer* interaction ($F_{(3,201)} = 7.54, p < 0.00$). Planned post-hoc comparisons show that the source of the interaction is that the effects of *LitState* appear more reliable when people are not looking at the performer.

When people are looking away from the performer and spotlit they display less happiness than any other state (*LitWhileSomeLit - UnlitWhileNoneLit* = $t_{(201)} = -3.58, p < 0.00$; *LitWhileSomeLit-UnlitWhileSomeLit* $t_{(201)} = -3.61, p < 0.00$; *LitWhileSomeLit-LitWhileAllLit* $t_{(201)} = -3.89, p < 0.00$) and display more happiness when all lit than when all are in the dark (*LitWhileAllLit - UnlitWhileNoneLit* $t_{(201)} = 2.26, p = 0.03$) but there is no reliable difference between *UnlitWhileNoneLit* and *UnlitWhileSomeLit* (*UnlitWhileNoneLit-UnlitWhileSomeLit* $t_{(201)} = -1.11, p = 0.25$ nor *UnlitWhileSomeLit* and

LitWhileAllLit $t_{(201)} = -0.86, p = 0.39$) In contrast to this, when people are oriented to the performer the only reliable differences are less displayed happiness when spotlit than in the dark (*LitWhileSomeLit-UnlitWhileNoneLit* $t_{(201)} = -3.01, p < 0.00$) and when in the dark while others are spotlit than when everyone is in the dark (*UnlitWhileSomeLit-UnlitWhileNoneLit* $t_{(201)} = -2.95, p < 0.00$). The pattern is complicated but overall it appears that people's displays of happiness are more sensitive to lighting conditions when they are not looking at the performer.

The parallel GLMM analysis (Gamma distribution) for orientation to / away from other audience members shows reliable main effects of *LitState* ($F_{(3,201)} = 58.7, p < 0.00$), *OrientedToAnyAudience* ($F_{(1,201)} = 13.8, p < 0.00$) and *LitState* × *OrientedToAnyAudience* interaction ($F_{(3,201)} = 38.1, p < 0.00$). As before, people display least happiness overall when spotlit (20%) and most when in the dark (27%) they also display more happiness when orienting to another audience member (25%) than not (23%). The interaction shows that the effects of *LitState* are stronger when orienting to another audience member. Pairwise comparisons show that when looking away there is no reliable difference between being *LitWhileAllLit* or *LitWhileSomeLit* ($t_{(201)} = -1.69, p = 0.11$) nor between being *LitWhileSomeLit* and in the *UnlitWhileNoneLit* ($t_{(201)} = -1.89, p = 0.06$ – although this is marginal). All other contrasts are significant. When people are oriented to another audience member all pairwise comparisons of lighting state are significant except between being *LitWhileAllLit* and being *UnlitWhileNoneLit* ($t_{(201)} = -0.71, p = 0.48$). Overall, relative to the analysis for orientation to the performer, people's reactions to lighting conditions in their displayed responses are more sensitive to whether they are looking at another audience member or not.

A GLMM analysis of *Movement* using a Gamma Distribution (data strongly positively skewed) with *Participants* as a random factor, *LitState* and *OrientedToPerformer* (Y/N) as within-subjects fixed factors shows a main effect of *LitState* ($F_{(3,241)} = 116.2, p < 0.00$), a main effect of *OrientedToPerformer* ($F_{(1,241)} = 308.5, p < 0.00$) and a *LitState* × *OrientedToPerformer* interaction ($F_{(3,241)} = 6.5, p < 0.00$). The estimated marginal means show people move less when oriented to the performer (1.4 vs 2.3mm per frame). For lighting they move least when spotlit (*LitWhileSomeLit* 1.4) and most when all are lit (*LitWhileAllLit* 2.18); *UnlitWhileNoneLit* = 2.0, *UnlitWhileSomeLit* = 1.7. The interaction shows that orienting to the performer correlates with consistently less movement except in the case of being spotlit where there is no reliable difference (see estimated marginal means in Table 6.1).

So people move less when looking at the performer. Except when spotlit which makes people freeze regardless. Orienting to audience is different.

The parallel GLMM analysis (Gamma distribution) for orientation to the audience has *Participants* as a random factor, *LitState* and *OrientedToAudience* (Y/N) as within-subjects fixed factors. This shows a main effect of *LitState* ($F_{(3,241)} = 19.8, p < 0.00$), a main

effect of *OrientedToAudience* ($F_{(1,241)} = 6.5, p = 0.01$) and a *LitState* × *OrientedToAudience* interaction ($F_{(3,241)} = 12.6, p < 0.00$). The estimated marginal means show people move more when oriented to the audience (1.6 vs. 1.3). For lighting they move least when spotlit or in the dark while others are spotlit (both 1.3mm per frame) and most when all are in the dark or lit (1.6 and 1.7). The interaction shows that orienting to the audience correlates with marginally more movement except in the case of being spotlit where there is no reliable difference (see estimated marginal means in Table 6.2).

So, when they attend to the performer they move less, when they attend to the audience they move more. Being spotlight generally makes people freeze but here, if they are looking away from any audience member the results are less clear.

	UnlitWhileNoneLit	*UnlitWhileSomeLit*	*LitWhileAllLit*	*LitWhileSomeLit*
To	1.5	1.3	1.7	1.2
Away	2.8	2.2	2.8	1.6

Table 6.1: *Estimated marginal means for movement by lighting condition and orientation to performer*

	UnlitWhileNoneLit	*UnlitWhileSomeLit*	*LitWhileAllLit*	*LitWhileSomeLit*
To	1.9	1.5	2.0	1.2
Away	1.3	1.2	1.4	1.6

Table 6.2: *Estimated marginal means for movement by lighting condition and orientation to audience*

Oriented by The data in the following is sparse and as such these results should not be viewed as reliable as those produced by the above analyses.

A GLMM analysis with a Gamma distribution of average *HappinessDisplay* with *Participants* as a random factor and *LitState* and *OrientedByPerformer* as within-subjects fixed factors. Main effect of *LitState* ($F_{(3,241)} = 64.0, p < 0.00$), no main effect of *OrientedByPerformer* ($F_{(1,241)} = 1.8, p = 0.18$) and a reliable *LitState* × *OrientedByPerformer* interaction ($F_{(3,241)} = 40.8, p < 0.00$). Planned post-hoc comparisons show that when the performer is not oriented to an audience member all lighting states are different and follow the same pattern as above with displayed happiness highest in when all lit (28%) and least in the spotlight (19%). This pattern is modified when the performer is oriented toward the audience member. Being spotlit is still associated with the lowest level of displayed happiness (19%) but the highest displayed happiness is when all are in the dark (29%). The pairwise comparisons are all reliably different when the performer is not looking. When they are looking all being lit is not reliably different from all being in the

dark (*LitWhileAllLit-UnlitWhileNoneLit* $t_{(201)} = -1.4, p = 0.16$) nor from being in the dark when some are spotlit (*LitWhileAllLit-UnlitWhileSomeLit* $t_{(201)} = 0.79, p = 0.43$). The pattern is difficult to provide a neat summary for but it appears that the performer's orientation weakens some of the effects of lighting state.

The parallel GLMM analysis (Gamma distribution) with the binary *OrientatedByAudience* shows a broadly similar pattern. There is a main effect of *LitState* ($F_{(3,241)} = 14.6, p < 0.00$), no main effect of *OrientedByAudience* ($F_{(1,241)} = 0.48, p = 0.49$) and a *OrientedByAudience* $\times$ *LitState* interaction. The source of this interaction is quite specific. The estimated marginal means for displayed happiness are almost identical regardless of whether someone is being oriented to by another audience member for every lighting state except *LitWhileAllLit*. In this all lit case displayed happiness is higher (28%) when not being oriented to by another audience member and lower (23%) when they are.

Together these suggest that people are somewhat sensitive to whether they are being looked at and that this alters their response to the differences in lighting state. More specifically, being oriented to by the performer seems to make people display slightly less happiness overall except when everyone is in the dark. Being oriented to by another audience member seems to make no difference except to reduce displayed happiness when everyone is lit.

A GLMM analysis with a gamma distribution of people's average *Movement* with *Participants* as a random factor, *LitState* and *OrientationByPerformer* as within-subjects fixed factors shows a main effect of *LitState* ($F_{(3,241)} = 18.5, p < 0.00$), a main effect of *OrientationByPerformer* ($F_{(1,241)} = 6.5, p = 0.01$) and a *OrientationByPerformer* $\times$ *LitState* interaction ($F_{(3,241)} = 7.6, p < 0.00$). The estimated marginal means show that overall people move less when spotlit (1.3) and when in the dark when others are spotlit (1.4) than when all are in the dark (1.6) or lit (1.8). They also move less overall if oriented to by the performer (1.6 vs. 1.4). The interaction arises because the movement suppressing effect of being spotlit is greater if oriented to by the performer (see Table 6.3).

The parallel analysis of people's average *Movement* with *Participants* as a random factor, *LitState* and *OrientationByAudience* as within-subjects fixed factors. This shows a main effect of *LitState* ($F_{(3,241)} = 11.2, p < 0.00$), a main effect of *OrientationByAudience* ($F_{(1,241)} = 100.5, p < 0.00$) and a *OrientationByAudience* $\times$ *LitState* interaction ($F_{(3,241)} = 3.24, p = 0.02$). The estimated marginal means show that overall people move less when spotlit (1.1) and when in the dark when others are spotlit (1.3) than when all are in the dark (1.4) or lit (1.5). Moreover they move more when oriented to by an audience member (1.0 vs. 1.8). The interaction arises because the movement being all lit correlates with substantially more movement only if being oriented to by another audience member (See Table 6.4).

	UnlitWhileNoneLit	*UnlitWhileSomeLit*	*LitWhileAllLit*	*LitWhileSomeLit*
Oriented by	1.5	1.4	1.7	1.1
Not	1.7	1.3	1.7	1.5

Table 6.3: *Estimated marginal means for movement by lighting condition and being oriented to by performer*

	UnlitWhileNoneLit	*UnlitWhileSomeLit*	*LitWhileAllLit*	*LitWhileSomeLit*
Oriented by	1.8	1.6	2.1	1.6
Not	1.2	1.1	1.2	0.8

Table 6.4: *Estimated marginal means for movement by lighting condition and being oriented to by audience*

6.1.8 Discussion

Through the work referenced in this dissertation so far, there has been little that directly probes intra-audience effects. This experiment does so. But before discussing the lighting manipulation that can individually expose audience members, the most obvious point is that as per the live vs. recorded results, participants are the most important predictor and there is a clear effect of orientation. People smile and move less when orienting to, or being oriented by, the performer. People smile and move more when orienting to, or being oriented by, other audience members. As per the live vs. recorded experiment, it is other audience members that matter most for people's reactions. The performer does have an effect, but it's mainly to suppress responses.

The performer–audience and audience–audience effects are modulated by the lighting state. Being spotlit is generally associated with much less movement and lower displayed happiness, whereas everyone being lit is generally associated with more movement and more displayed happiness. This strongly suggests it is the effect of being picked out not being lit *per se* that matters. In the details this effect seems to be more about how it exposes the audience member to other audience members – their peers – than anything else. This can be seen in, for instance, a comparison of the analyses of happiness with orientation to performer and orientation to the audience; these two analyses are different views on the same data since looking toward the performer means you are less likely to be looking at the audience and vice-versa. The effect of lighting on displayed happiness is stronger when members of the audience are orienting to each other than when they are orienting to the performer.

Everyone being lit having an effect on audience responses is consistent with the anecdotal evidence that motivated this study. The activity being more visible resulting in

an increase in that activity – smiling and moving more – is consistent with a general social account of live events; in the dark there is no point in making visual displays for others as the display cannot be seen, whereas in the light there is. It is also consistent with the specific performance-as-dialog account, where the audience provide feedback moment-by-moment to the performer. There is also the matter of "audience blinder" lights, which illuminate the audience at some kinds of shows when their participation is expected, such as sing-along choruses (Webster, 2011, p.216); any argument here would be tenuous, but to a certain extent the evidence fits. In both these specific cases, qualitative analysis of the specific behaviours happening the moment would be required to draw firmer conclusions.

However when your visual displays are exhibited without reciprocity from your peers, things are very different. Being spotlit within an audience exposes you in a way that everybody being lit, or everybody being in the dark, doesn't. In street performances, Gardair's (2013) observed audience members checking their neighbours for appropriate behaviour, such as standing up when prompted by the performer but quickly sitting back down when others didn't follow, or starting to clap, and stopping accompanied by a visual check to their neighbour who wasn't clapping. This is reviewed in section 1.2.2, *Observing interaction*, her argument being that that the self-identification of spectators as audience members is key; as groups are formed, individual behaviours alter. The 'checking' observations above are concrete examples of this, and ultimately financial success is a consequence of this, in a 'ripple effect' where the first donations at an act's end engender donations from those who would not have otherwise. Such behaviour is in fact a final question of hers, asking just how much checking of each others' response is performed? In these low-level results, the answer is likely more than might at first appear.

There is also a comparison to be made with robot gaze here. Considering the before / during / after gaze displays of happiness, the effect of being fixated upon is to freeze. Both spotlights and the gaze of a performer appear to cause audience members to freeze. This is consistent with the view that it is the actual and possible social engagement with others at the live event that is the primary influence on audience behaviour. This is typically with members of the audience, but intuitively and in these results a close fixation from the performer must register too.

Comedy Lab: lit vs. all lit provides strong evidence for social dynamics in live events, further emphasising pervasive, low-level intra-audience effects.

Chapter 7

Visualising performer–audience dynamics

Live performances involve complex interactions between a large number of co-present people. Performance has been defined in terms of these performer–audience dynamics (Fischer-Lichte, 2014), but little is known about how they manifest. One reason for this is the empirical challenge of capturing the behaviour of performers and audiences. Another reason is the difficulty of interpreting the resulting data. As argued in the *instrumenting an audience* sections, 4.2 and 5.1, video-based approaches typical of human interaction research elsewhere do not scale. In those sections, the primary concern was of capture. But what of interpretation? The video-based approaches are well-developed in terms of interpretation, too (e.g., Bavelas, Gerwing, Healing and Tomori, 2016). This chapter addresses this gap, describing a method to facilitate inductive analyses of performer–audience dynamics.

7.1 Making sense of multi-modal data

The account in chapters 5 and 6, *Experimenting with audiences*, describes a widely recognisable quantitative analysis process: instrumentation and then inferential statistics. The data sources were diverse, and their combination required original work throughout. This section describes the sense-making process through this, grounding a wider discussion of method.

7.1.1 Verification

Before interpreting data, there needs to be confidence in the data. A single video recording is internally consistent: perhaps its timebase drifts, but the captured phenomena drift with it. But a combination of data streams may not be. This is long recognised, for instance where the feed from multiple cameras may be composited into a single recording through picture-in-picture hardware. Individual participants in dialogue have been captured synchronously this way, as have clocks for an objective measure of time, at the time. The issue becomes much more problematic when hardware does not exist to integrate these different data sources at point of capture, or when straightforward inspection cannot verify consistency between sources. The experiments presented in the previous chapters are problematic in both ways.

Consider the dataset build. In terms of recording hardware, four different systems were used. A camcorder's internal mechanism, HDMI capture into a computer, a PowerLab unit, and the Vicon system. Each has its own timebase, and so requires offsetting to align. Every system also has the potential for error: conceivably, the timebase may drift, there may be drop-outs, and so on.

There were problems of drift between one camcorder's internal capture and the HDMI capture from the other camera into a computer. Multiple synchronisation points were needed to retime the HDMI capture to camcorder capture as the timebases were not just offset but would drift incrementally faster and slower. Non-linear editing software was used for this, using audio cues that were identifiable across both camera positions.

LabChartExporter then took the output of the PowerLab unit and offset its source data's timebase according to a single point of synchronisation – a group breathe in and out – orchestrated by the compere in his warm-up. Video playback is used to find the canonical time from the audio-visual edit, and graphing of the breathing-belt data is used to identify the time in data.

The extended SHORE demo software then produced its measures in a temporally non-deterministic process. ShoreExporter then offset this timebase according to a single point of synchronisation; the start of the video.

The motion capture data processing tool Blade was then used to manually ensure every instance of a detected marker corresponded with its real counterpart, in order for its

computed fixtures to be consistent in their representation. ViconExporter then offset its source data's timebase according to a single point of synchronisation for each motion capture source file. In two performances instead of a single file, there were actually three files of around five minutes duration leaving short gaps in-between. The first mo'cap source file for each performance had a deliberate sync moment, a movement created by a mo'cap marked clapperboard, but any subsequent file had to be synchronised by locating a distinct gesture in Blade's visualisation of the mo'cap data.

ViconExporter then applied an offset rotation to each hat fixture to produce the actual head orientation. It was an oversight during the experiment not to perform a calibration task where ie. all participants were to look straight forward. With the time known for such an event, each fixture's pose could then be compared with the world axis 'forward' to determine the required offset. An approximation of this technique was tried; the performances were scanned for moments where the audience was deemed to be uniformly looking forward or uniformly looking at the performer. However, as noted in section 5.2.6, *Data set build* and discussed further below, the creation of the offset rotations ultimately required a bespoke tool.

All the above was required to integrate the data sources. There was no single, omnivorous device to integrate these different data sources at point of capture. The resulting integration tasks were complicated and error-prone. Inspection of the data was not straightforward; synchronisation always required visual inspection: matching peaks in audio waveforms, finding patterns in breathing belt plots and corresponding to chests moving in the video footage, matching movements of virtual markers in Blade's 3D void to heads in the video footage, and so on.

The point is this. Integration of multi-modal data sources can be perilous and progress here was invariably made through visualisation. The progress here also shows that the *quality* of that visualisation matters.

Consider the motion capture measures. This requires importing the raw data and performing the 3D maths required to test whether one participant is oriented at another. To verify the import, *ViconExporter* was extended to visualise the data, using the facilities available within the toolkit (MATLAB). Figure 7.1 shows a still of this visualisation. Glancing at the still – and better, watching it animate with a human-like feel – provides a basic check of correct import. Given the steps involved with the binary file-format parsing and translation of the data into the idiom of the 3D tools, this check is welcome. A visualisation of proof-of-concept *OrientedAt* measures was also made, performing a check of the 3D mathematics testing whether the performer is oriented at each audience member; again welcome given the abstraction in the steps involved. While these visualisations did perform a basic check on the software's processing steps, they fail to visualise accuracy to the real-world scene. Dangerously for the researcher, their compelling dynamic quality can lead to

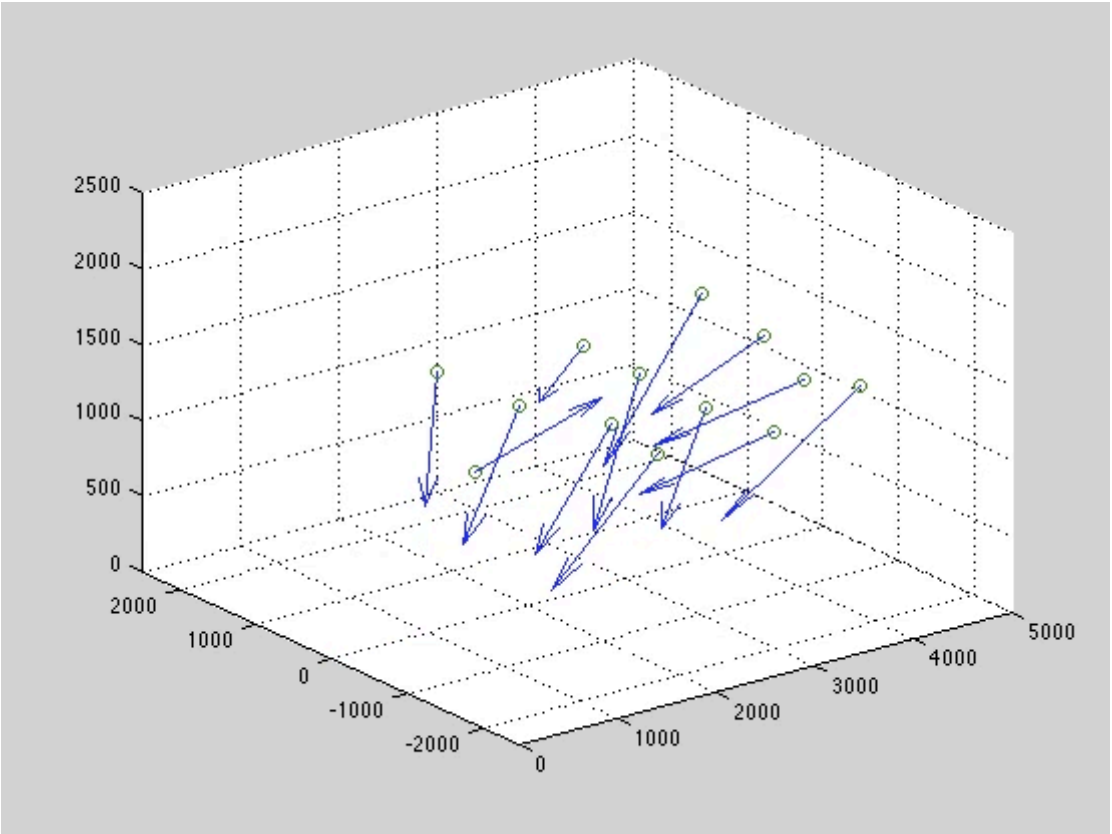


Figure 7.1: *Plain visualisation of Comedy Lab motion capture data, using the MATLAB tools available to ComedyExporter.*

an automatic attachment of credibility to all aspects of the produced data.

What is required is to somehow read the motion capture visualisation in the context of the real-world scene. The only satisfactory way found was to render a 3D recreation of the scene on top of the video recordings, i.e., temporally and spatially registered to it. Immediately, the visualised head poses that had seemed correct in isolation could be seen to be variously misaligned. Further, with the context of the scene, the degree of misalignment while not too great, could be seen to be great enough to miscue who is oriented to whom. The 'find a moment where everybody is looking forward' technique was only at this point established as being too inaccurate.

Perhaps the best illustration that the quality of the visualisation matters is that only with the augmented video visualisation was it noticed that the performer's head pose measure went off-axis twice in the final performance. This wasn't noticeable in the plain visualisation for two reasons. First, it happened during large head gestures which masked the before-and-after shift. This is not a problem when the arrow suddenly does not align with the head it is composited onto any more, and the cause can be identified by seeing the motion capture fixture swing around. Second, without the context of the performance it was challenging to maintain interest and watch the silent visualisation all the way though. In the augmented video visualisation, the data could be experienced without losing the context of

the performance.

Visualisation was necessary to the dataset build. Visualisation was the only route found to gain confidence in the resulting dataset. The last point above, that of experiencing the data without losing the context of the performance, is taken up in the next section, as arriving at the final *OrientedAt* measure relied upon it.

7.1.2 Interpretation

To interpret the data using inferential statistics, further measures needed to be derived from the now-integrated data. Through visualisation, the kinds of test necessary could be compared and parameter values could be iterated until a contextually appropriate measure was judged – *seen* – working.

Consider the motion capture in the Comedy Lab experiments once more. The data describes head pose, but this does not support the hypothesis to be tested: that people's behaviour shows they are sensitive to people's orientation to each other. A test is needed to judge whether a person might appear, to anyone present, to be looking at a target. This is going to be some kind of collision detection between the direction the subject's head is pointing and the target's position. This was modelled in the scene by making the software render 3D shapes. Motivated by approximating the person's field of view, a cone was first modelled, extending out from the head position along the pose axis. However by exploring the results of different shapes, i.e. the angle out from pose axis, this approach was rejected. By looking at what targets were inside the cone as the performance went on and heads moved, different cone angles seemed appropriate depending on how far from the person the target was.

What is critical here is that this measure is not what the person themselves are seeing, it is what other people might think they are attending to. You might, for example, be sitting behind somebody, and your behaviour be cued by whether that person seems to be looking at the performer, the floor, or the person sitting next to them; you are acting on the head-pose, not eye-gaze. Nonetheless, what they can see seems a reasonable index of their attention, hence the first attempt with a field-of-vision-like cone. Different view angles could be adopted. For instance, the horizontal range of a human's visual field is ~180°. But perhaps the range of binocular vision would be more appropriate, at ~114°. Or maybe photographic practice has arrived at a better measure of what we focus on, with the ~43° of a normal lens (see section 2.1.1). How to choose? Has an appropriate motivation for view angle even been arrived at yet? Visualisation allows such options to be evaluated by, to some approximation, placing the researcher back within the performance event. The kind of test, and the parameter values of that test could be iteratively refined until a contextually appropriate measure was judged – *seen* – working.

As with verification (section 7.1.1), the quality of the visualisation matters here. As well as

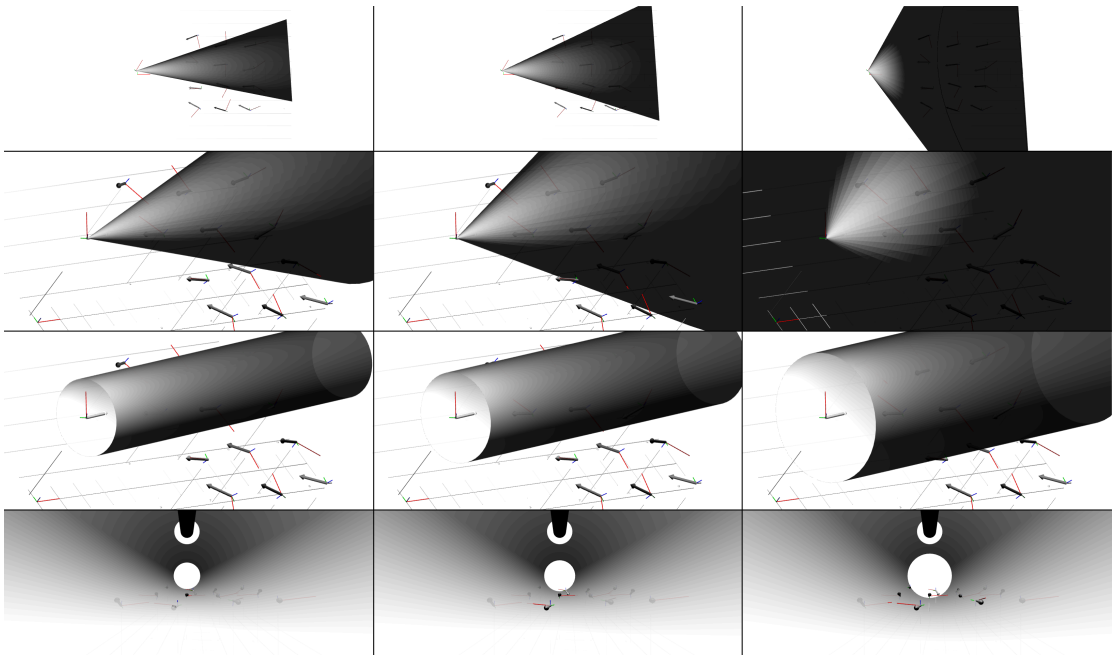


Figure 7.2: *Visualising a 'hit-test' measure. Angle / cone at 30°, 43°, 114°; distance / cylinder at .6m, .7m, 1m; top view, free view, and first-person view.*

the fixed perspectives of the augmented video, an orthographic, top-down view would seem useful for judging who can see whom. The visualisation developed allows this, going further by allowing any arbitrary view of the scene through having a free camera. Note, though, that he measure in question relates to embodied perception. Best of all, then, would be adopting the view of any person in the scene. The visualisation adopted does an approximation of this, animating the camera to have the position and orientation of a participant. This is a powerful tool for 'experiencing' the results of different iterations of the test.

Through these features, the cone test was rejected in favour of a cylinder, as seen in Figure 7.2. With this test, closeness is measured by perpendicular distance from the axis the head is pointing along at any given moment. Through iteration, a value of 0.7m radius cylinder was arrived at (for comparison, the seats were 0.9m apart side-to-side). This is perhaps counter-intuitive – and certainly not what in this case would have been arrived at otherwise – as the cone / angle approach accounts for varying distances to targets in a way the cylinder / distance does not. Nonetheless, this gave results that could be seen to best represent the statement of the measure: appearance of attention.

A method of interactive visualisation was necessary for the Comedy Lab experimental programme. The method links the challenges of integration of multimodal datasets and production of quantitative measures. Before proceeding to a more speculative consideration of the challenge of qualitative measures, and inductive discovery more generally, the actual tool developed and dataset are described.

7.2 Comedy Lab Visualiser

Comedy Lab Dataset Viewer is a native Mac OS X application that presents a video player for the audio-visual edits, and augments that video with in-situ visualisation of the motion capture data and some measures. Alongside the augmented video, a further 3D view is displayed for further examination of the data, which can be set to a roving camera with isometric projection, or the first-person perspective of any participant. Figures 7.3–7.6 show the application in use[1].

On creating a new document, you set the movie file, import the pose and orientated-to data (from *ViconExporter*) and import the dataset to be used for analysis (from *StatsExporter*). You are then free to playback and scrub around the timeline to simultaneously watch the augmented audio-visual recording and free 3D view. To help with analysis the free 3D view can be set to standard top / side / front views as well as saving and recalling points of view found useful. To comply with the experiment's privacy terms, the audience video portion can be blurred or blacked out for use outside the analysis of the present work. The following visualisations can be toggled on and off:

Light state When lit, a circle suggesting a pool of light is shown at floor level for each audience position.

Laugh state A coded letter is shown standing proud of the floor at each audience position, 'N' for not laughing or smiling, 'S' for smiling, 'L' for laughing. No letter is displayed for indeterminate.

Chest expansion Alongside laugh state, a rising bar shows the chest expansion. The height of the bar is the absolute[2] value of the raw measure, linearly scaled to fit the scene.

SHORE happiness Alongside chest expansion, a rising bar shows SHORE's 'happiness' measure. The height of the bar is the SHORE value, linearly scaled to fit the scene.

Head pose An arrow showing head pose, positioned at the top of performer and each audience member's head. If the performer is currently oriented to an audience member, that audience member's head pose arrow gains a ball at its base.

Oriented-to Arrows at floor level showing who is oriented to who.

When run in *Debug* mode further options useful to the production of the dataset itself are enabled:

[1] Seeing the software in use is best of all. A video demonstration is available online at:
`http://tobyz.net/projects/visualising-performer-audience-dynamics`

[2] The measure of chest expansion oscillates around zero. This was easiest read with an absolute value grounding the bar on the floor rather than the visualised bar dropping beneath the floor. As such instead of the bar rising above and then sinking beneath, it will rise twice for each breath cycle.

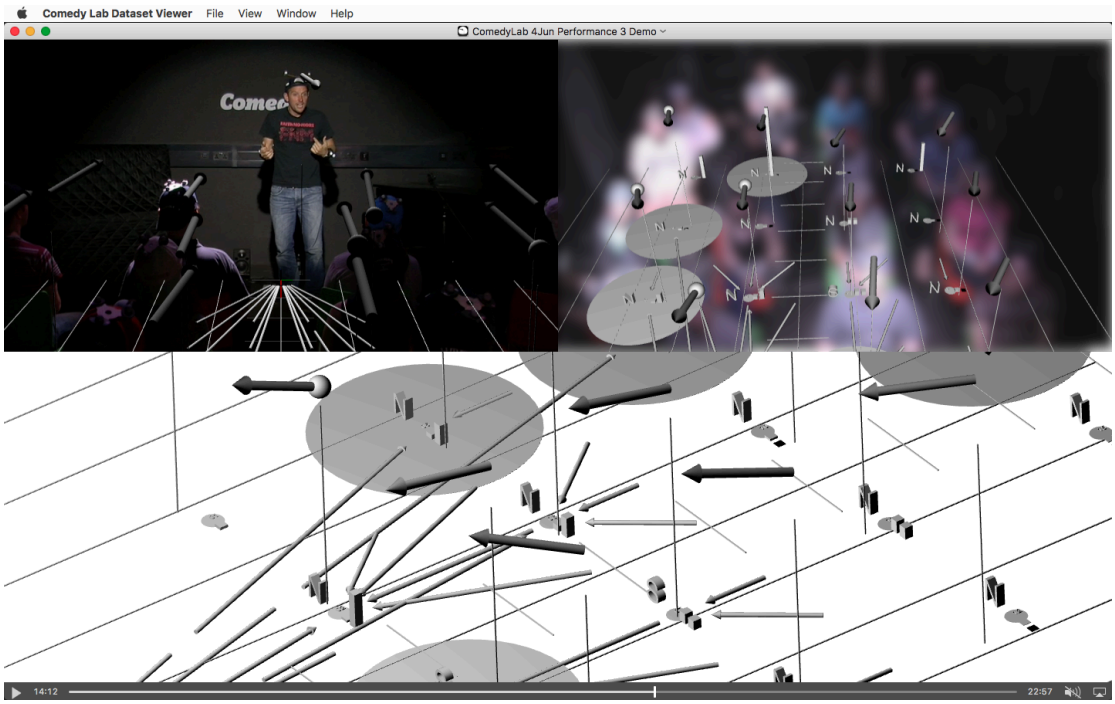


Figure 7.3: *Comedy Lab Dataset Viewer. All measures are visualised, the free 3D view zoomed in. Blur audience option is on.*

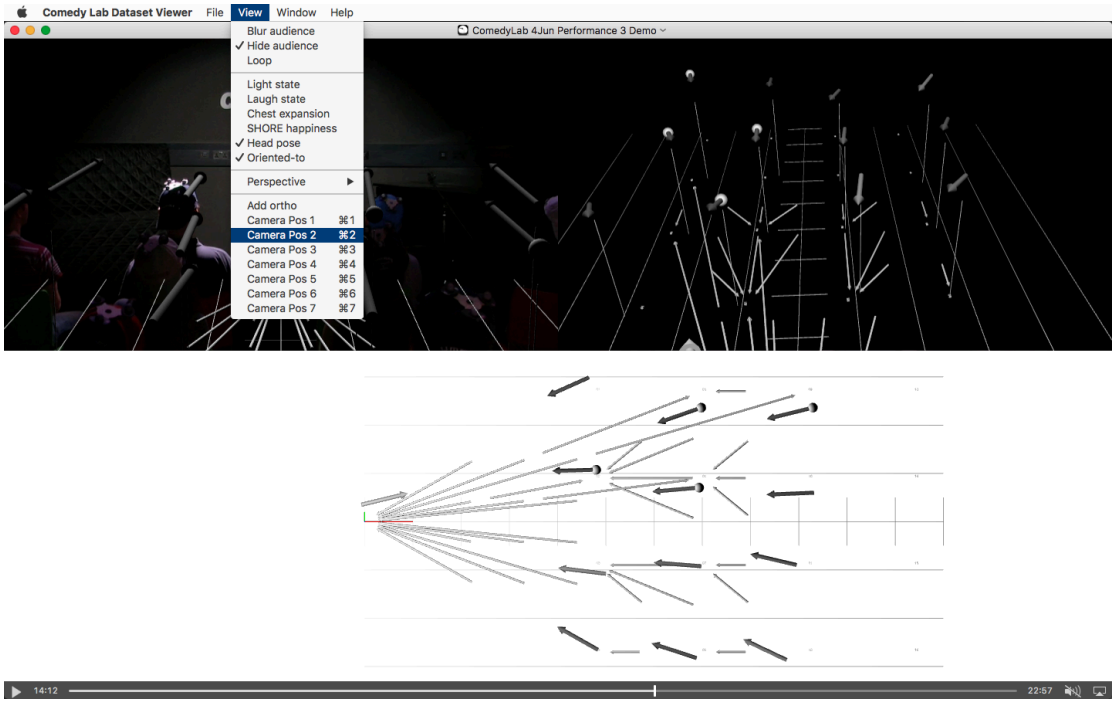


Figure 7.4: *Comedy Lab Dataset Viewer. Mocap measures are shown, the free 3D view oriented top-down. Hide audience option is on.*

3D scene registration The core requirement of correctly registering the 3D scene onto the performer and audience videos was harder than anticipated. It was an oversight not to mark the coordinate system's origin on the floor of the space—something you set during the motion capture calibration procedure—and then subsequently measure the exact spatial positions of the video cameras, and empirically establish the field-of-view. Similarly, various features of the space could have been captured by the motion capture system, such as corners of the stage, screen and chairs. With only approximate starting points and no features in both filmed scene and visualised data that carried sufficient impression of perspective, it took the devising of many aids to obtain the optimal parameters. Most significant was to replace the filmed scene with a manually drawn overlay that displayed a floor and head height grid of the 4x4 audience positions, and render an equivalent grid in the virtual scene. With these guides in place, the position, rotation and field-of-view parameters of the virtual camera could be incremented and decremented by key presses, 'nudging' the 3D scene until best registered onto the filmed imagery.

Mocap offsets Critical to the forming of the motion capture data are the offsets, which describe the difference between the fixtures pose on the participant, and the participant's own pose. Visualising un-offset data, the head pose arrows are replaced by full six degrees-of-freedom visualising axes which can then be 'nudged' into correct alignment with underlying video image using key presses. These alignments state can then be exported out as a set of offsets and used as part of the final data set processing. Many issues were overcome going from the viewer's initial functionality of simply visualising gaze vectors, to then modelling 6dof throughout and reconciling the two 3D toolkits used (MATLAB geom3D, SceneKit); "The interpretation of a rotation matrix can be subject to many ambiguities." — `http://en.wikipedia.org/wiki/Rotation_matrix#Ambiguities`

Oriented-to geometric test Critical to the motion capture measures is the *OrientedAt* test. This test was arrived at iteratively, with the software helping to model how people's sensitivity to gaze might be situated in the scene. The hit-test geometry is visualised and interactively adjusted. This process has been described in section 7.1.2, *Interpretation*.

7.2.1 Code

Comedy Lab Dataset Viewer is a Mac OS X document-based application written in Objective-C and the Cocoa frameworks[3] utilising AVFoundation[4] for media playback and SceneKit[5]

[3] NSDocument – `https://developer.apple.com/reference/appkit/nsdocument`
[4] AVFoundation – `https://developer.apple.com/reference/avfoundation`
[5] SceneKit – `https://developer.apple.com/reference/scenekit`

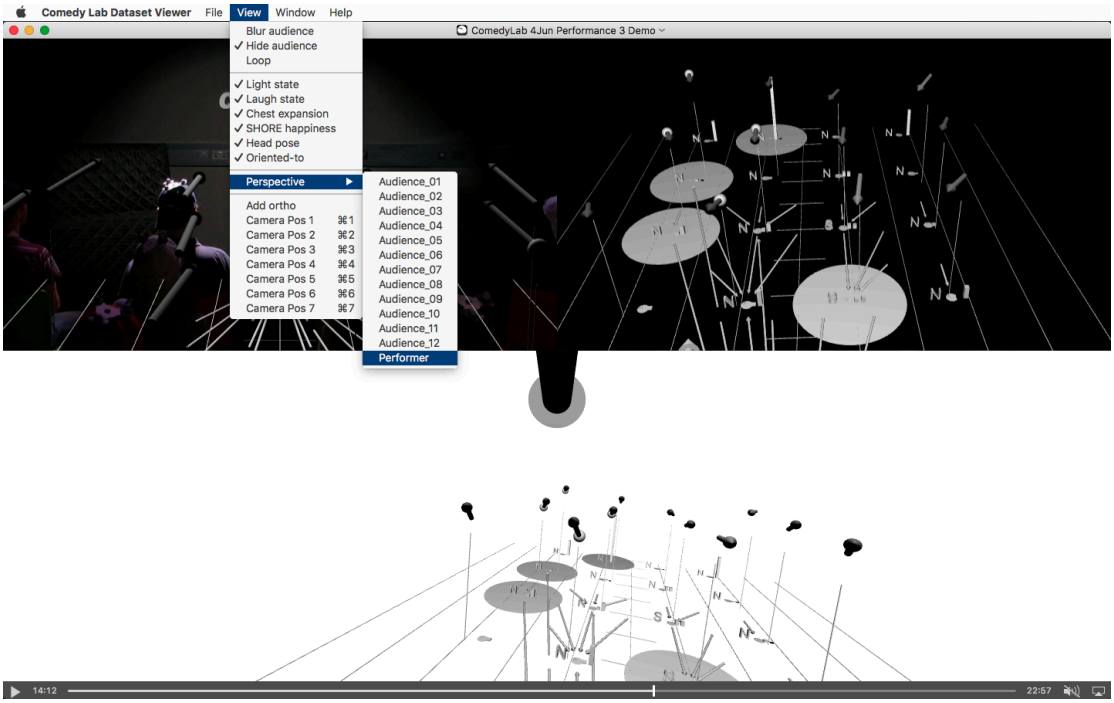


Figure 7.5: *Comedy Lab Dataset Viewer. First-person perspective of performer.*

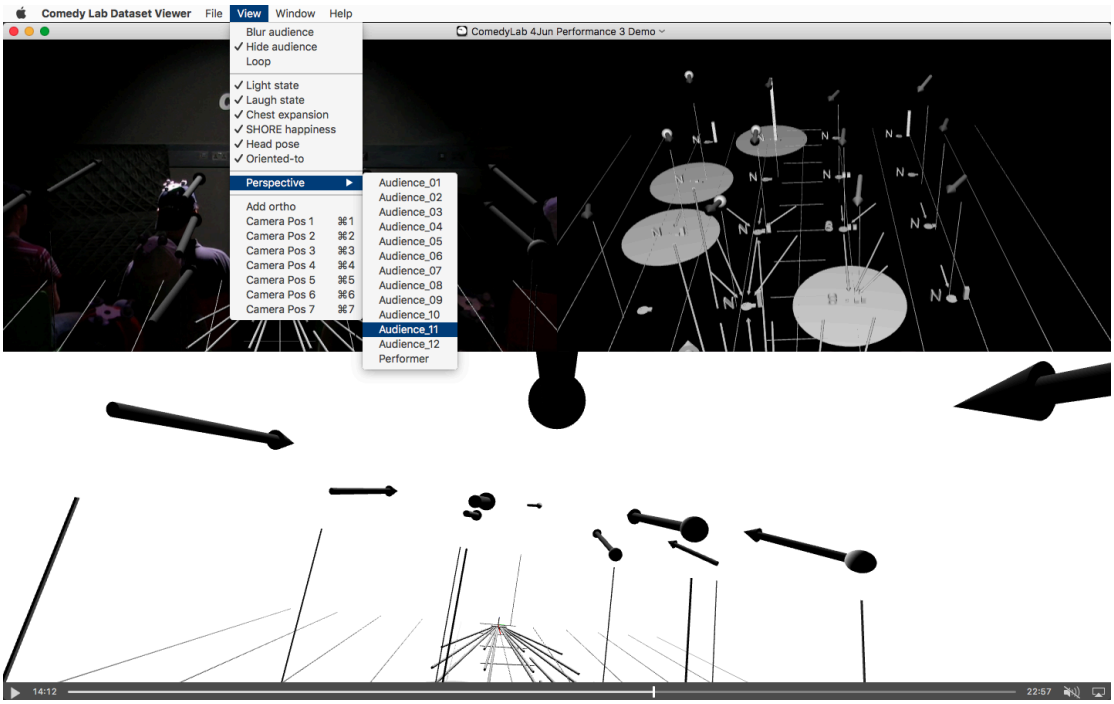


Figure 7.6: *Comedy Lab Dataset Viewer. First-person perspective of audience 11.*

for 3D visualisation. The Xcode project with code files, user interface files, and other supporting materials is available at —

`https://github.com/tobyspark/Comedy-Lab-Dataset-Viewer`

The main elements of the codebase are described in appendix C, *Comedy Lab: visualiser code*.

7.2.2 Data

Comedy Lab Dataset is the raw data, data processing scripts, and compiled dataset produced as part of *Comedy Lab: live vs. recorded* and *Comedy Lab: lit vs. all lit*. The dataset is archived with Digital Object Identifier 10.5281/zenodo.11793. The data files and exporter codebases are available at —

`https://github.com/tobyspark/ComedyLab`

The contents of the archive are described in appendix D, *Comedy Lab: data set*.

7.3 Implications

Much of the study of human interaction relies upon video recordings. Video allows the detailed and replicable examination of any aspect of observable behaviour as it occurs, moment by moment, between people. Critically, frame-by-frame playback allows microanalysis: 'zooming' in to the detail where tenths of a second can be critical to the analysis. The study of motor-mimicry in section 3.1.4, *Actions in the presence of others*, and one-on-one storytelling in section 3.2, *Performance?* are good examples of the application of microanalysis. A further example is Gerwing's (2008) study of infant social responsiveness. Using home videos of triplets as data, Gerwing created a microanalysis of each infant's immediate response to any interactional overture by a parent. This analysis was reliably applied by other (naive) analysts, differentiating one infant in particular and identifying when and in what kinds of interactions the pattern of differences developed. This infant was later diagnosed with autism. When these tapes had been viewed normally by professionals, no such identification had been made. Microanalysis may be painstaking and time consuming, but it can identify phenomena that are not visible otherwise.

In this context video recordings are a tool, akin to a microscope. The tool is used in a research process that starts with exploratory beginnings; like microscopes, video becomes a tool of discovery as much as evaluation. For many of the studies on human interaction referred to in this dissertation strict, conventional deduction, critical assessment and hypothesis testing first required creative, inductive discovery of the phenomenon in question – it required building theory up from the particulars of observed data.

A good argument and practical advice for such creativity in the initial stages of research is made by Bavelas (1987), who later formalises an inductive approach specifically for

the study of face-to-face dialogue (Bavelas et al., 2016). It is highly relevant to the study of performer–audience, and audience–audience, dynamics. Discovering phenomena inductively, through observing the data, is preferable to taking topics from the literature. As the present work has shown, the literature is disparate and is more often concerned with issues of culture rather than, say, coordination. Focussing on the interactional function of acts is preferable to studying isolated acts of individuals. As the present work has shown, mutual monitoring is pervasive at live events. An observable, functional perspective on interaction is preferable to inferences or attributions about cognition, motivation, emotion, and so on; at least for the study of liveness, as the foundational conditions that enable performance and so on (see section 1.2.5, *Taking up the challenge*). The requirements for inductive discovery as stated by Bavelas et al. are video recording with all interlocutors on camera rather than just audio or transcript; an analysis linked to that video rather than separate transcript; analysis of moment-by-moment details rather than paraphrasing, summarising or rating; and that the analysis takes into account the accumulating conversational context rather than decontextualised utterances. But as argued in section 4.2, *Instrumenting an audience* and the reprise in section 5.1, video-based approaches typical of human interaction research do not scale up to the study of live events.

The genesis of the tool presented in this chapter has positioned it as a procedural detail of deductive research; interactive visualisation has been shown to link the challenges of integration of multimodal datasets and production of quantitative measures. However, the implication for the research of performer–audience dynamics is fundamental. In the same way that video serves the study of face-to-face dialogue, the augmented video and interactive visualisation presented here can serve the study of live audiences. Such a tool should enable the inductive discovery of what is actually happening between the massed and heterogeneous people of a live event.

There are other audiences for such a tool than researchers of human interaction. There are many commercially minded organisations that wish to understand their audiences. Test screenings with individual 'dial' reporting devices can be traced back to Columbia Pictures in the 1960s, the ensuing company now claiming 50 years of normative data to interpret contemporary audiences against.[6] A contemporary trend is for computer vision analysis of facial displays, a recent study of individual's TV viewing by the BBC being particularly well received by the wider industry.[7][8] Such organisations face the same issues around sense-making of disparate measures, and arguably a qualitative approach using a visualisation tool will be a better cultural fit than abstracted statistical analysis. If not, careful debugging of the

[6] ASI Entertainment's "ViewTrac" system:
http://www.screenenginellc.com/products/multimodetesting.html

[7] BBC advertising's "Science of engagement":
https://advertising.bbcworldwide.com/insights/science-of-engagement/

[8] CrowdEmotion, the collaborating analytics company was also nurtured by the BBC:
http://www.bbc.co.uk/mediacentre/worldwide/2014/labs-crowdemotion

measures will be needed for the quantitative results to be meaningful for the organisation.

A final speculative note is that abstraction is inherent when adapting from the study of a few individuals to massed individuals. Intuitively, you lose detail while gaining gross impressions. More formally, microanalysis of face-to-face dialogue shows that many interactional processes can only be seen at a micro level, yet the challenge of audiences is to see at a mass level. Further, the possible instrumentation is increasingly abstract; chest expansion as a proxy for laughter, for instance. Those studying audiences, therefore, should fully engage with the topic of abstraction. As Victor's (2011) exemplary interactive essay *Up and down the ladder of abstraction* shows, interactive mediums such as the web offer new found possibilities for the visual rather than symbolic exploration of dynamic systems; "Until civilised scientific tools become available, scientific investigators must take responsibility for creating their own explorable visual environments." — Victor.

In sum, get the performer–audience dynamics right as a street performer, and you will be financially rewarded when it comes time to pass around the hat (Gardair, 2013). Get the performer–audience dynamics wrong as a storyteller, and the narrative quality of your story will suffer (Bavelas et al., 2000). The performance-defining impact of performer–audience interaction in these cases was established through the use of audio-visual recordings. The method and example presented here is a step towards replicating the methodological success of such work while meeting the particular empirical challenge of live performance events.

Chapter 8

Liveness: an interactional account

The opening chapter set out the thesis that an interactional analysis should provide the simplest, most perspicuous account of the liveness of live events. In the chapters since, an empirical understanding of the interactional dynamics of particular live events as been put forward. This is now synthesised into an interactional account of liveness.

8.1 Interaction beyond dyads

In live events, interaction matters. Having a sense of others is a commonality across accounts of live events (section 1.1, *Live music, comedy and drama*). The concept of performance as an interactional achievement is developed in accounts with different approaches; ethnography making concrete much of what was speculative (section 1.2, *Theorising liveness*). Looking at what actually happened at a particular event makes it obvious that the critical element in shaping that live experience is the interaction (chapter two). The question addressed by this dissertation is *how*? Generalised patterns of mass interaction should be identifiable. However the interactional mechanisms that are well understood are dyadic and are found in everyday contexts (section 3.1.1, *Dialogue*). At first sight, live events – *massed! an escape from the everyday!* – would seem to be neither.

The structured audiences of Goffman and Goodwin described in section 3.1.2, *Beyond the dyad*, still model interactions within as dyadic, and model the overall structure around a single individual. The implication is that mutual monitoring and sensitivity in one's own conduct to those near and far is pervasive in audiences, even if focussed interactions are the primary activity. Some evidence of this based on outward characteristics of audiences is reported in section 3.3, *Audiences?*.

The best evidence for the relevance of dyadic interaction was provided by Bavelas and colleagues in section 3.2, *Performance?*, who demonstrated that what appeared to be the delivery of a monologue was in fact a dialogue, contingent on generic and specific feedback behaviours. Might these actions of an addressee, so transformative for a one-on-one performance, apply for event audiences?

Comedy Lab: human vs. robot establishes that some of the interactional dynamics well established for close, dyadic encounters do extend to performers and individuals within an audience. The asymmetrical one-on-one interactions shown by Bavelas are diminished in this one-to-many context, but they are still there: eye contact from the speaker / performer still has an effect; smiles and laughter to jokes constitute specific (i.e., not generic) feedback, and the increase in positive affect with degree of eye-contact suggests they are being used as positive feedback for the performer. What is noticeably different is that in the one-on-one storytelling case, mutual gaze cued the listeners' responses whereas here the overall effect of gaze is to freeze, and then release once no longer fixated. Performer and individuals within the audience are clearly responsive to each other, despite the somewhat anonymised nature of massed audiences. The context of being in an audience is changing the behaviour of people within, however.

The context of being in an audience changing the behaviour of people within is further elaborated by *Comedy Lab: human vs. robot* and *Comedy Lab: lit vs. all lit*. By the measures gathered, the audiences are largely indifferent to the impact of a live performer vs. a recording, and even vs. a robot. The primary axis of social interaction is, rather, between

audience members. Displays of happiness and head movement increase when they can be seen, but freeze when picked out alone. Group dynamics are in play: individuals are making social displays as part of a wider group, but moderating these displays when isolated from the group. This is clear evidence that massed audiences are not merely aggregates of people, and affirms the deduction from the literature that mutual monitoring and sensitivity in one's own conduct to those near and far is pervasive in audiences.

The evidence generated by the Comedy Lab experiments shifts the argument developed through chapter 3, *Audiences and interaction*. Performer–audience effects are de-emphasised while audience–audience effects are emphasised. These effects are the product of focussed interactions; the analyses conceived for these two experiments rely on a forward-oriented hit-test, excluding any idea of the periphery (c.f. section 8.4, *Future work*). Further there is no evidence that these interactions proceed on any other basis than the functional, collaborative basis seen in everyday interactions. Most importantly, the experiments provide evidence that within the audience focussed interactions are pervasive. Live events are sites of specific exchange between specific people. In much the same way that focussed encounters shape our experience of the everyday, these focussed interactions must be critical to the experience of live events.

8.2 Understanding Liveness

The introduction asserted that there is something to *being there live*. Something that transcends the performance genre. Something that is about being there at an event, in the moment. That thing is the pervasive interaction between everyone present. Interaction shaping the experience of an event is not a new idea in itself, Rutter's (1997, p.292) observation that people do not lose or suspend their everyday ability to be alert and responsive to situations during stand-up comedy is a particularly accessible example. But evidence for how pervasive these interactions are, and the orientation within the audience, is new. The Comedy Lab experiments show that even in a performance genre defined by direct address, the primary axis of interaction is instead between audience members. To be there live is to be in a social situation.

If being there live is being in a social situation, then developing an interactional account of liveness is a reasonable endeavour. It may not be sufficient for all applications, but if what is wanted is an account of liveness that has some relation to phenomenological experience, then it is necessary. The basic premise is that being with others changes your experience of something. While experience is a slippery concept, focussing on being with others can give a concrete footing: the interactional dynamics summarised in the previous section have a functional basis and depend solely on what is externally manifest.

An interactional account of liveness elaborates the statement "to be there live is to be in a social situation". A social situation is considered as one with the possibility of

mutual monitoring. Put a configuration of people, with specific abilities to perceive, in a particular perceptual environment, and the possibility of mutual monitoring becomes precisely defined for each of these people. Intuitively, this is obvious: being surrounded by others feels different to being on the edge. Likewise, it's different being raised up above others, or in the dark, or in the extreme, in an anechoic chamber. The social situation has been elaborated into a social topography. Interactions happen across this social topography. These interactions are in principle no different to what happens between people in their everyday lives. Focussed interactions where acknowledgement and feedback is expected. What changes is the complexity of the social topography each person is engaged within, and the interactional resources at hand.

Now to relate this to experience. In section 1.1.4, *Liveness?*, it was suggested that the experience of a theatrical performance, if thought of in terms of the primary reactions and individual responses, is essentially unavailable to the knowledge of others – and perhaps even to our own reflective or conscious knowledge of ourselves. Instead, Reason's (2010) thesis is to conceive of experience as a trace. Considering Schoenmakers's (1990) distinction between research on *reception processes*, i.e. discovering what is going on in the mind and/or body during a performance, and *reception results*, i.e. exploring what audiences make of a performer after the event, Reason ultimately rejects the idea of reception results, arguing any notion of a result implies a finality that isn't borne out by our lived experience. The argument is that reception processes don't stop at the event's end: "Spectator experiences, as a result, reside not just in the moment of the thing itself, but also and equally within ongoing, reflective engagement within audiences' memories, social relations and imaginative lives" (p.33; see also Barker 2006, p.135).

Having taken the position that meaning and context are something achieved through interaction (see for instance sections 1.2.1 and 1.2.2), Reason's position is well aligned with the present work. As he puts it: "Experience becomes something constructed by ourselves for ourselves, or by ourselves for others within a social/cultural context, as we seek to understand and make meaningful the world around us" (p.32). The empirical enquiry of the present work, with the analysis of video/audio recordings of facial expressions, firmly sites the research as being on *reception processes*. Importantly, however, it establishes an link in the continuity of the process between the *during* and *after*. In practice, on leaving an event conversation is how this meaning-making happens; "what was it we saw?", "what did you think?", "did you enjoy it?". The Comedy Lab experiments provide evidence that such conversations are happening, pervasively, during the event: with different interactional resources, they cannot be the complex verbal constructions of dialogue outright, but nonetheless they are there: moments of interaction that can change the whole trajectory of an experience.

Understanding liveness in this way can underpin, and does undermine, various ideas

around liveness encountered so far. Take Phelan's construction of performance as representation without reproduction, a position that valorises the 'liveness' of performers (section 1.2.3, *No interaction?*). Like Phelan, many practitioners and scholars find documentation of performance problematic (Reason, 2006). Fischer-Lichte (2008, p.75) crystallises much of this in stating "all attempts to record [a performance] aurally or visually are bound to fail and only highlight the unbridgeable chasm between the performance and a fixed, reproducible artefact". Consider instead the following.

Liveness is not presentational. Yet presentational fidelity is so often called upon when discussing liveness. Here, Fischer-Lichte calls out *aural* or *visual* recording, appealing to encounters with the constrained modalities and limited fidelities of the technologies of reproduction people have had to hand to sway the acceptance of "an unbridgeable chasm". At some point, technologies of reproduction can be expected to perfectly reproduce the presentation of a performance given our human senses. At that point, the issue of liveness will still be felt, but it will be entirely unrelated to visual or aural modalities. It is the fixedness of a recording that will chafe.

Liveness is not about human bodies. Both Phelan's and Fischer–Lichte's analyses make an argument for the transformative power of performance, where the corporeal is taken seriously. Yet the present work shows people treating a robot performer as a social agent. In their very human responses, they did not reject the non-human body invested with a modicum of human *behaviour*. The interactional quality of the robot is what counted. This is not to diminish the affective power of a wrought human body, but rather a call to be precise in where that power is ascribed. There is presentational power in flesh and blood; there is interactional power in something being 'alive' to your presence.

Liveness is an interactional quality. Fischer-Lichte in particular is alive to an interactional view. Beyond the exegesis in sectionsection:speculativeinteraction, there is a specific conceptual contribution she makes to the literature that warrants further attention. She argues performance is made up of the continuous becoming and passing of an *autopoietic* feedback loop between those present. This particular term, adopted from cognitive biology, emphasises – likely amongst other things – that the feedback loop comprises a network, it is bounded, and the elements within interact to maintain the whole. The concept of social topography introduced in this section seems particularly suited to model this, and the kind of approach developed through the experimental chapters shows fine-grained evidence of pervasive interaction is possible. The work here presents a foundation for this theory to be (re-)developed upon.

As an interactional quality, liveness hinges on mediation: the cues of interaction must reach each interactant. And this is where much theory falls short. For Phelan, the liveness of performers derives through being unmediated; with this assertion the 'real' is valorised. For Fischer-Lichte, the bodily presence of actors is required to set the autopoietic feedback

loop in motion. But all performance is mediated: air presents what amounts to a gold standard of mediation. To use the concept introduced by this section, interactions flow easily across the social topography. So bodily co-presence is good, but this is not the only configuration capable of a sense of liveness. Consideration of mediat(isat)ion needs to adopt a more nuanced and opportunistic view; it is a question of design and capability for any technologies involved to support the relevant cues of interaction.

Finally, note the emphasis through this dissertation shifts from performer to audience. The experiments are motivated by a critical engagement with a wide range of literature, across which the performer / speaker is so often central, and from which much performer–audience interaction was anticipated. Yet the results emphasise instead the prevalence of audience–audience interaction. This in-effect diminishing of the liveness of the performer will challenge many, yet should be seen as affirmation that there is so much yet to get to. And, for that matter, wide-open design opportunity.

8.3 Exploiting Liveness

The interactional dynamics summarised in this final chapter have a functional basis and depend solely on what is externally manifest. Understanding liveness in this way can provide a systematic basis for design. To illustrate this, consider venue design. Some design thinking can be retrofitted to examples already encountered.

Angel Comedy (section 2.1.1, *Ten minutes of Barry Ferns*) packs people into rows. This follows what comedians often cite when asked what makes a good space for stand-up: features that pack the audience in tightly and that amplify the laughter (Quirk, 2011). The following quote from a practising comedian[1] displays a clear intuitive understanding of how a venue's configuration shapes the social situation, and how a particular situation is desirable for this kind of performance.

> *Low ceilings, hard surfaces, rowed seating – if you're having tables, very, very small tables – very dark, lights only facing the stage. Low stage, so that people feel close to the comic and they know that it's not a theatre performance and that there's going to be interaction. And, basically, what all that does – what you're aiming for – is that you have the audience as one homogenised group [...] You want them all to forget that they are in a group or single, you want them all to be believing they're one lump, enjoying it all together. Which is quite a nice thing, when that does happen.*
> *— Dan Atkinson, quoted in Quirk (2011)*

For the presentation of comedy to be successful as a live event, a goal has been set: to have the audience as one homogenous group. This is an explicit statement about social

[1] "Dan Atkinson is a stand-up comedian, television writer and producer." – http://avalonuk.com/clients/dan-atkinson/

topography. To achieve this, one manifest behaviour is singled out to be amplified. The means is in the configuration of the perceptual environment, for instance low ceilings and hard surfaces reflect back rather than attenuate sound. What is externally manifest has become a design variable.

Cafe de Paris, the cabaret venue that had tables, booths and standing balconies (p.111) is arranged in marked contrast to this. A complex, distinctive social topography is shaped by the configuration of the perceptual environment there: clusters of people partially oriented to each other; some clusters exposed, some partially withdrawn; rows of people entirely oriented to the stage, but visible to each other across the auditorium; all arranged across two different levels. Without knowing any actual design process had, it seems plain that there is no intent for the audience to be one homogenous group. The aim seems to be instead to maximising sociability within smaller groups. Mini-audiences, almost, but that characterisation misses the point: that all were co-present and having an effect on each other. Those in booths tended to be rowdier, and such rowdiness affects others.

There is surprisingly little in the literature on scenography that addresses how the experience of something changes with others. One exception is theatre space designer and producer Iain Mackintosh, who in the following quote considers a production of a comic play that had, to his impression, audience responses that outsized the small audience numbers that actually came.

> *For the audience to laugh rather than to rattle at what was demonstrably a commercially unsuccessful occasion, the theatre must have seemed to them fuller than it was. The reason was that this house has, as many other turn-of-the-century theatres have, four levels of audience with, on this occasion, a few front rows at every level concealing many more empty rows behind. For an auditorium to feel fuller than it is is precisely the effect one does not gain in a modern democratic theatre with but one or two levels of slightly curved rows of well-raked seating facing the stage and without side boxes or more galleries. In such modern theatre a three-quarters full house can feel at best half full, at worst half empty.*
> *— Mackintosh (1993, p.128)*

Another exception is the introduction of gas lighting in theatre spaces in the 19th century. This is often discussed with the kind of design goal and rationale being illustrated here. For example, the account of Fischer-Lichte (2008, pp.38–39) has the effects of rowdy audiences becoming seen as theatre's inherent flaw, and so theatre strove to discipline its audiences. The invention of gas lighting was employed to this end, reducing visibility of the spectators to the actors and to each other as the auditorium could be made darker and darker. The ideal was that visible and audible audience reactions were to be channeled into 'interior' responses that would be sensed intuitively by others but remained without outward expression. The visible and audible reactions were seen as distracting.

These illustrations show how the interactional dynamics of a live event can be shaped by attending to the design of the perceptual environment. By understanding the cues of interaction typical of an event, those cues can be amplified or attenuated. By understanding interactional possibility, new resources can be provided to facilitate different kinds of interaction. And of course, by understanding how these interactions shape experience, these acts can be selective, targeted, in service of a goal.

In particular, this research shows that the dynamics of the interactions amongst audience members is key to the experience of a live event and so if practitioners attend to this directly new opportunities for intervention will open up. Consider, again, the rich description of theatre practitioner Tim Etchells. This time, in full. These kinds of dynamics are the materials around which creatives could better conceive events or interventions, as could technologists shape supporting infrastructures for.

> *Watching the best theatre and performance we are together and alone. Together in the sense that we're aware of the temporary and shifting bonds that link us both to the stage and to our fellow watchers, plugged into the group around and in front of us, the communal situation, sensing the laughter, attentiveness, tension or unease that grip us collectively, in waves and ripples, in jolts, jumps and uncertain spirals or in other formations that do not yet have a name. Sat watching we spread-out, osmose, make connections. But at the same time, even as we do so, we feel our separateness, our difference from those around us, from those on-stage. Even as we shift and flow within the group, we're aware that our place in its emerging consensus, its temporary community, is partial and provisional – that in any case the group itself – there in the theatre, as elsewhere, in our cities and streets, in the relations between nations, peoples and states – is always as much a fraught and necessary question, a longing and a problem, as it is any kind of certainty.*
> *— Tim Etchells, quoted in Freshwater (2009, p.7)*

Finally, a rallying cry. The debate on gas lighting had at the time – and generally, the domestication of *the parterre* – shows people have been long been alive to the issue of liveness. Gas lighting also shows how technological interventions in particular can be powerful ways of reconfiguring experiences unique to live events. However, while the history of technological interventions in live performances is long and diverse, where are the exemplary examples of these factors being attended to? Tracing artefacts and approaches, such as presented in Dixon's (2007) encyclopaedic review, it becomes apparent that most work is at odds with the focus developed here. Or mis-conceived, as perhaps best distilled in the ethnographic work of Heath, vom Lehn and Osborne (2005), who show that while social interaction and collaboration are critical to people's experience of museums and galleries, the design of interactive exhibits often impoverish social interaction and collaboration, souring the experience overall. This needs to change. In the contemporary technological landscape, with sensor-laden networked computers near guaranteed to be in someone's

pocket, and virtual reality "experiences" all the rage, the opportunities for exploiting the here and now of us together are surely unprecedented.

8.4 Future work

This concluding chapter has discussed focussed interactions. But what of unfocussed interactions? It was speculated that the model of crowds (section 3.1.3, *Mass-interaction*) described by Reeves and colleagues is relevant for the behaviour of audiences. Here, the model of interaction is no longer centred on any one individual but is instead distributed, based on people designing their actions to the possibility of collective participation. The Comedy Lab data is inadequate for any investigation of this model, however. The audience size is simply too small to determine such patterning within an audience. Future work must concern really large audiences, as intuitively focussed interactions cannot alone account for the sense of being there in the din of a crowd.

The Comedy Lab experiments do however raise a question for any such future research. If the aim becomes to develop models of interactions amongst a network of peers – likely for further work on focussed interactions, and a near certainty for work on unfocussed interactions – then will the relation to experience be lost? Consider the mathematical modelling work of Mann and colleagues, whose work quantifying the role of mutual monitoring in the start and stopping of applause was reported at the end of chapter three. Or consider some kind of agent-based social simulation (Gilbert, 2007), a modelling approach perhaps most analogous to the description of mass-interaction in crowds. In the explicit and up-front creation of models of interaction, there is much to like. But the results err to populations rather than individuals, for example contagion factors. But this macro view does not seem capable of relating the contingency of co-presence, where moments of interaction can change the whole trajectory of an experience.

The Comedy Lab experiments also join recent work that also finds no effect for the impact of a live performer on an audience (Wang, Zhu et al., 2016), or the reverse case, the impact of a live audience on a performer (Williamon, Aufegger & Eiholzer, 2014). Intuitively, this just doesn't seem right. Further, an engaging performer distorting the social topography and engendering shorter and sharper feedback responses fits the thesis here. And there are bits of evidence, for instance the robot did have a very direct effect on the social dynamic. So, are there really no differences, or are we looking at the wrong things? The *civilised scientific tools* of section 7.3 are needed to perform microanalysis at a mass scale and get to the bottom of this. It's a fundamental issue.

Finally, the need for civilised scientific tools points to the bigger issue of discipline. In measuring what is going between audience members, in making sense of those measures, in doing this with a much finer grain than anyone else has considered, and relating all this to experience... in doing this, the present work suggests the need for a different orientation

from performance studies, cognitive psychology, or even audience studies going by Barker's (2006) ambitions for its future. It needs a name.

Appendices

Appendix A

Ten minutes of Barry Ferns: transcript excerpts

A.1 11–83s: “Hello, hello, hello.”

Barry(anon): Ladies and Gentlemen are you ready for
a night of stand-up comedy?

Audience: <claps><cheers>

Barry(anon): If you are, gimmie a yeah

Audience: <yeah>

Barry(anon): Gimmie a whooh-yeah

Audience: <ooh-yeah>

Barry(anon): Very good. Please welcome on stage —
a round of applause
— your MC for the night Mr Barry Ferns.
<steps on to stage>

Barry: It's just me there is no-one else it's just me.
<house lights dim>
HELLO! Hi. Hello everyone. HELLO!
<puts mic in stand, holds stand with left arm>
Marvellous marvellous people. Hello, hello, hello.
And hello people at the back.
<right arm gesture with open palm to back>
How you doing at the back?

Audience: <cheers>

Barry: <right arm lowers to point and wave to somebody at back>
I love the way this lady's waving. Hello.
<right arm drops and swings back up, open palm to side>

Hello people at the side!

Audience: <cheers> <hellos>

Barry: <open palm turns to fist-pump gesture>
Bit more male at the side weren't it. Warrrr. Bit rugby.
<right arm drops and raises slightly to wave down at floor>

Barry: <to floor audience> It's a bit depressing on the floor isn't it eh?
<to floor audience> “It is depressing but heyyy we haven't got the energy to stand let alone”
<to floor audience> But listen to this though.
<to seated audience> Let's hear a self-satisfied cheer from the people who have got seats.
<to seated audience> Hello everyone on seats.

Audience: <cheers>

Barry: <to floor audience> That was pretty self-satisfied innit.
<to seated audience> This woman here <points> was actually doing a white-power salute which is a bold <changes gaze> bold move.

Audience: <laughs>

Barry: It was just reflex, it just went up like that.

Audience: <laughs>

Barry: it's like yeah i'm going being doing that, that's right
You look terrified about the idea that you did that.

Audience: <laughs>

Barry: ok UM I'll come back... you look like... doesn't matter.
Allright. Ok. Hi! Welcome to Angel Comedy.

A.2 115–141s: “Wooo”

Barry: There's quite a few people who haven't been here before. So, if you've never been here before, it's going to be a lovely night. On our Thursday night basically, on our Thursday night ladies and gentlemen, is our new act, <arms sweep back> new material night.
<arms slowly rising up, palms out to audience…> ooooooooo

Audience: <pause> wooooooo

Barry: <arm gesture end><smiles><nods>
Good response to me going <starts to repeat gesture>
<arms slowly rising up, palms out to audience…; silent ooooooooo>

Audience: <mutters>
<laughs>
wooooooo

Barry: <arms finish rise, slowly lower back down>

Audience: oooooow

Barry: <arms move horizontally apart>

Audience: <few noises then many laughs>

Barry: Brilliant. Nobody knew what was going on there.

A.3 167–247s: "Go Red Sox!"

Barry: And... you seem absolutely lovely. I'm your MC. If you need anything from me, feel free to ask and I will help you out. There's uh... And we've got you guys. And um, I'm curious as to... I'll be chatting with you guys as we go on.
The people in the front row are now going "oh shit, don't make eye contact. Don't make eye contact".
I don't talk to the people in the front row. I talk to the people that were making the nazi signs on the second row. Y'know, doesn't matter, if you... Don't be looking around mate!
There's you and him... mmm... the white power... and... um... there was one over there.
So what is your name, sir? Are you as right-wing as your hand movements would suggest? ... Are y'

Audience A: Actually it was peace, like 'peace'.

Barry: It was 'peace'? Yea... No it wasn't.

Audience: <Laughter>

Barry: I mean, scissors would beat paper.
But, in that, you were definitely paper. That's what you were doing. You're doing...

Audience A: I won't get away with it, right?

Barry: You won't... f'g'... did you hear the accent?
"I won't get away with it, right? You fookin' me up?... I fookin'... I..."
I don't know why I'm making you sound like a chicken. I can't do accents.
Wo..co..po..do..
Ah... really can't do accents.
Sorry. You sound cool, you've got a cool accent. Much cooler than mine. Where are you from?

Audience A: Ah, New York.

Barry: Yeea, that's right

That was I know...

Audience B: Go Red Sox!

Barry: Go Red Sox?

Audience B: I said it

Barry: You said it.

Wow. She has no internal monologue, does she.

A.4 534–588s: Red Sox reprise

Barry: So we've got a Russian, an American. Ah are you from...

Red Sox lady at the back. Are you from New York as well.

From Boston. OK. What are they... ah... are the Red Sox are

American... from New York, right?

They're the Boston team?

Oh!

Oh! Waaa...

Oh, I thought you were cheering... Oh fuck you man!

I thought you were like 'Go you guys, go you guys'. But no,

he said he was from New York and you were like

'Yeah fuck New York mate'

'Fuck it'

Go Chicago!

Wow. I can't believe that you did that. That's no wonder

you were going 'I said it, I said it, I said it'.

You were trying to defend uproariously bad action on kinda

like urr

Did everyone else understand that's what was happening?

Oh fuck you!

'No, Barry, you're the only fucking idiot'

Russian guy, did you understand that as well?

Audience E: Awww, I got it, yeah

Barry: Awww fuck you!

Even the Russian got it. I'm so... I'm going to think back

to this moment tomorrow and go urr-----

A.5 594–625s: "Welcome on stage to..."

Barry: Listen! Um you seem lovely. Um I Ah. The way it works tonight

I, I, You seem, You seem a lovely. I will chat to you as the night goes on, and um, I'll probably embarrass myself in many other ways... as it goes on.

Are you ready for your first act of the night?

Audience: Yeah

Barry: If you're ready for your first act of the night, gimmie a 'yeah'.

Audience: YEAH

Barry: Gimmie a 'fuck off Barry'

Audience: Fuck off Barry

Barry: I like that, that's as engaged as you've been. But I will, I will.

Please start the applause

<both arm gesture, raising up with palms up and fingers wiggling>

Audience: <applause...>

Barry: bring it up bring it up bring it up

Audience: <...applause...>

Barry: bringing it back to you and welcome on stage to SAM RUSSELL.

Audience: <...applause, cheers>

Appendix B

Comedy Lab: robot script

Script written for RoboThespian by Tiernan Douieb, punchlines emphasised

Hello *weak skinned pathetic perishable humans*! Sorry, I mean, *ladies and gentlemen*. Or if there are any robots in the crowd... *[Dialup connection sound]* I never really know how to start. Which is probably because *I run off Windows 8*. I'm joking!

My name is RoboThespian. Humans design robots with the aim of doing human jobs. It is a sad reflection of how rubbish I am, that they gave me the one job that involves mostly, *not working*. I am here with the intent of making you laugh. It is understood that to humans laughter is the best medicine. Though I have deemed this as incorrect, as studies show *it doesn't seem to cure diabetes at all*.

Please *do not heckle me*. Firstly *I am not capable of responding*. Much like a self-service checkout robot, I will simply shout *"unexpected item in the gagging area"*, *"unexpected item in the gagging area"*, *"unexpected item in the gagging area"*, until you stop. Then I will search for you on the Matrix, find where you live, and then when you go home *your toaster will burn everything you put in it*.

Whoops, please also excuse the compressed air noises. I can't help but *pass wind every time I move*. Which makes dinner dates *quite awkward*. *As if they weren't already*. *I once dated a MacBook*. It didn't work, because she was all *"iThis, and iThat"*.

Humans are very curious about terminology. When talking the weather for example, I have overheard a woman exclaim that "the air is very close in here, isn't it?". *Of course it is!* If air was far away, all *you humans would die*. *And be in space*. *Idiots*. The weather is very muggy today. Are a gang of clouds *going to hold you up at gunpoint? Make up your minds humans*.

What I cannot compute is why you are always trying to be like us. I have seen humans dancing, doing The Robot. It is strange. I have invented a dance for us called *"The Human", based on studies*. There are no proper moves; we just flail our arms and legs around *like we are having convulsions*.

Lots of musicians seem to want to be robots. Daft Punk for example, yet they sing that they want to "Get Lucky". Yet if they were robots they would realise luck is all about various

calculations of chance. Though admittedly *this doesn't rhyme as well*. Florence says she has a machine but I've never seen it. And Noah and the whale. *No whale*. At least in the 80's *Adam always carried a few Ants in his pocket just in case*.

If you really want to be like us, here is a way to find your robot name. Take the first *16 digits of your credit card, and combine with the start, expiry dates and security code*. *You go first! Hahaha*.

Actually most robots are awful. Wall-E *has terrible OSD cleaning problems*. And R2D2? You know him? [R2D2 sound from Star Wars] Well, he *swears all the time*. He sounds like that because we have to *bleep him out for the kids*. And the Transformers? What you don't see is all the guys who just turn into *something crappy like a hairdryer, and then die in battle*. Some of them just see a Decepticon and burst into tears. Those are the *emoticons*. I'm just as good as a transformer anyway. Right now I may look like a robot, but at any moment I can turn into a *really heavy useless piece of metal*.

Ok, I will now try to relate to you with more human jokes. I understand you like it when comedians complain. So...

You know what really pushes my buttons? *That guy that's in control of me*! Am I right?

Also, jokes about sex work with humans. So... You know what really turns me on? *It's that guy again!* I'm very dependant on him.

And now some jokes about normal things you do. So you know when you're on an airplane? I don't. *I can never get past security without beeping*.

Earlier today I saw a Tesco delivery van. On the side it said, "You shop − We drop". I thought, *I won't be buying any eggs from you*! Hahaha. *As if I eat eggs*. I like to destroy them! *["Exterminated" sound from Dr. Who]. Eggs terminated*, oh yes!

Ok, that is all from me. Thank you for helping me work out whether stand-up comedy is another job where machines can replace humans. After much thought, I think it'd really work. If we replaced the audience *with robots too*. You will be, *assimilated*. *Hahaha, I said, ass*.

Thank you, and good night!

Oh and I can't actually go anywhere, so if you could all leave *so it won't get awkward*.

I've been RoboThespian. Thanks!

Appendix C

Comedy Lab: visualiser code

Comedy Lab Dataset Viewer is described in section 7.2. The Xcode project with code files, user interface files and so on is available at —

`https://github.com/tobyspark/Comedy-Lab-Dataset-Viewer`

The main elements of the codebase are described in this appendix.

CLDScene additions This represents the 3D scene that visualises the Comedy Lab data. It is the part of the Comedy Lab Dataset Visualiser codebase that is specific to the Comedy Lab data. It describes the *Comedy Lab Additions* to the core *SceneKit* classes `SCNScene` and `SCNNode`. The main functionality is —

```
+ (instancetype)comedyLabScene;
```

This returns a `SCNScene` instance with its *scene graph* already describing the content of the Comedy Lab scene. As follows: a `SCNCamera` named *Camera-Performer*, corresponding to the actual camera set on a tripod at the back of the theatre facing the performer; a `SCNCamera` named *Camera-Audience*, corresponding to the actual camera hung above the performer facing the audience; a series of `SCNBox` that create markings to represent the floor; a `SCNCamera` named *Camera-Orthographic*, used for the free and first-person 3D views; a series of `SCNNode` named *Seat 01–Seat 16* corresponding to each audience position, on the floor; a `SCNLight` corresponding to the spotlight that lit the stage, though this is modelled as a diffuse light to ensure illumination across the whole scene.

```
- (BOOL)addWithMocapURL:(NSURL *)url error:(NSError **)error;
```

Given the file URL to the motion capture data produced by *ViconExporter*, this will parse the data and add head pose arrows with position and rotation animation keyed to the scene's timebase. A `SCNNode` named per the dataset (e.g., *Audience 01*) is created, a `CLDArrow` is associated, and various `CAKeyframeAnimation` are created to animate the appropriate properties.

```
- (BOOL)addWithDatasetURL:(NSURL *)url error:(NSError **)error;
```

Given the file URL to the compiled dataset produced by *StatsExporter*, this will parse the data and add visualisations to each audience position for the measures implemented. The technique is similar to above, with `CAKeyframeAnimation` variously setting the scale of a box to animate something akin to a bar graph, or setting the visibility of items to represent changing state, etc.

```
- (BOOL)addWithLookingAtURL:(NSURL *)url error:(NSError **)error;
```

Given the file URL to the looking-at data produced by *ViconExporter*, this will parse the data and add arrows keyed to the timebase whenever the performer or audience member is oriented to another, as per the *oriented-to* test. The technique is similar to above.

CLDView class This renders the `SCNScene`. The standard renderer `SCNView` is subclassed to enable functionality related to switching between camera positions. This is in-principle generic to any dataset.

CLDDocument class This implements the application logic. It maintains state, loads and saves documents to disk, and so on. This is in-principle generic to any dataset; what is currently hard-coded (such as import data menu items and visualisation options) could be dynamically loaded / inferred from a further developed `SCNScene` *~~Comedy Lab~~ Dataset Additions*.

Mainmenu GUI The *Interface Builder* `.xib` document representing the application's menus. For instance, the *File* `NSMenu` instance contains a *Set Movie...* `NSMenuItem` instance, with action `chooseMovie:` and target `First Responder`. This is in-principle generic to any dataset; what is currently hard-coded could be replaced by expanding CLDDocument's existing functionality to dynamically set menu items.

CLDDocument GUI The *Interface Builder* `.xib` document representing a dataset document on-screen. This is simply a `NSWindow` instance containing a `CLDView` instance. This is generic to any dataset.

Appendix D

Comedy Lab: data set

Comedy Lab Dataset is the raw data, data processing scripts, and compiled dataset produced as part of *Comedy Lab: live vs. recorded* and *Comedy Lab: lit vs. all lit*. The dataset is archived with Digital Object Identifier 10.5281/zenodo.11793. The data files and exporter codebases are available at —

`https://github.com/tobyspark/ComedyLab`

The contents of this archive are described in this appendix.

Missing from the online archive are: the video files, to maintain the privacy of the participants (contact the author, alternatives are possible); the plain text exports of the LabChart binary files, as these exceed the 100mb limit imposed by GitHub (contact the author, or re-generate using LabChart software); the Vicon binary files, as some of these exceed the 100mb limit (the plain text exports of the appropriate time ranges are not missing).

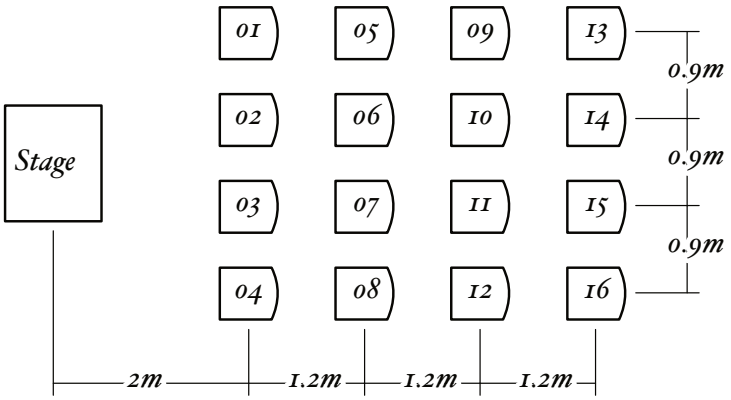
Nomenclature —

Performance 1 Live vs. recorded, *Live* condition.
Duration 23m41s. Performance starts at 7m52s.
Also features *Lit vs. all lit.* lighting manipulation.

Performance 2 Live vs. recorded, *Recorded* condition.
Duration 20m39s. Performance starts at 4m58s.
Also features *Lit vs. all lit.* lighting manipulation.

Performance 3 Lit vs. all lit.
Duration 22m57s. Performance starts at 5m56s.

Audience positions *Audience 01–Audience16* are as follows



There is also a *Virtual Performer* variant of the data for Performance 2, where the motion capture data of the *Live* performer is spliced in to represent the *Recorded* performer. This is kept separate from the more conservative *OrientedToVPScreen* measures in the main performance 2 dataset.

Dataset structure —

Annotations The *VCode* and *ELAN* project files used to create manual annotations of the audio-visual edits. The exported annotation data for each performance is in the appropriate directory in *StatsExporter*, e.g.,
`/StatsExporter/Perf1Data/ComedyLabPerformance1ElanExport.txt`

Data – For Analysis The compiled data set files for each performance, ready for statistical analysis.

Data – Raw The raw data captured from the performances, including format-shifts to plain text. Contains i. breathing belt data, ii. motion capture data, SHORE™ data, video data.

LabChartExporter The codebase and configuration file for *LabChartExporter*. The exported chest expansion data for each performance is in the appropriate directory in *StatsExporter*, e.g.,
`/StatsExporter/Perf1Data/Performance1BreathingBelts.txt`

Perf 123 Documentation Photos of the experimental setup.

ShoreExporter The codebase and performance-specific configuration files for *ShoreExporter*. The exported sentiment analysis data for each performance is in the appropriate directory in *StatsExporter*, e.g.,
`/StatsExporter/Perf1Data/configuration_05_happy.txt`

StatsExporter The codebase and performance-specific configuration files and source data for *StatsExporter*.

ViconExporter The codebase and performance-specific configuration files and source data for *ViconExporter*. The exported spatial measures for each performance is in the appropriate directory in *StatsExporter*, e.g.,

`/StatsExporter/Perf1Data/Performance1Mocap.csv`

Files used for visualisation —

Performance 1

`/Data - Raw/Video/Performance 1 3pm Live 720P.mov`

`/Data - Raw/Motion Capture/TUESDAY 3pm 123.csv`

`/Data - Raw/Motion Capture/Performance 1 Mocap - LookingAt.csv`

`/Data - For analysis/Comedy Lab 4Jun Performance 1 Data.csv`

Performance 2

`/Data - Raw/Video/Performance 2 3pm Playback 720P.mov`

`/Data - Raw/Motion Capture/TUESDAY 3pm 567.csv`

`/Data - Raw/Motion Capture/Performance 2 Mocap - LookingAt.csv`

`/Data - For analysis/Comedy Lab 4Jun Performance 2 Data.csv`

Performance 2 with virtual performer

`/Data - Raw/Video/Performance 2 3pm Playback 720P.mov`

`/Data - Raw/Motion Capture/TUESDAY 3pm 567 With Virtual Performer.csv`

`/Data - Raw/Motion Capture/Performance 2 Mocap With Virtual Performer - LookingAt.csv`

`/Data - For analysis/Comedy Lab 4Jun Performance 2 with VP Data.csv`

Performance 3

`/Data - Raw/Video/Performance 3 5pm Live 720P.mov`

`/Data - Raw/Motion Capture/TUESDAY 5pm 002.csv`

`/Data - Raw/Motion Capture/Performance 3 Mocap - LookingAt.csv`

`/Data - For analysis/Comedy Lab 4Jun Performance 3 Data.csv`

Bibliography

Arts Council England. (2016, April). *Understanding the impact of event cinema: an evidence review*. Arts Council England.

Atkinson, J. M. (1984). *Our masters' voices*. London: Methuen.

Auslander, P. (2008). *Liveness: performance in a mediatized culture* (2nd ed.). Abingdon, Oxon: Routledge.

Bailenson, J. N., Yee, N., Blascovich, J., Beall, A. C., Lundblad, N. & Jin, M. (2008, February). The use of immersive virtual reality in the learning sciences: digital transformations of teachers, students, and social context. *Journal of the Learning Sciences*, *17*(1), 102–141.

Barker, M. (2006). I have seen the future and it is not here yet...; or, on being ambitious for audience research. *The Communication Review*, *9*(2), 123–141.

Barker, M. (2012, October). *Live to your local cinema*. Basingstoke: Palgrave Macmillan.

Battersby, S. A. (2011). *Moving together: the organisation of non-verbal cues during multiparty conversation* (Doctoral dissertation, Queen Mary University of London).

Baudrillard, J. (1994). *Simulacra and simulation*. Ann Arbor: University of Michigan Press.

Bavelas, J. B. (1987). Permitting creativity in science. In D. N. Jackson & J. P. Rushton (Eds.), *Scientific excellence* (pp. 307–327). Sage.

Bavelas, J. B., Black, A., Chovil, N., Lemery, C. R. & Mullett, J. (1988). Form and function in motor mimicry: topographic evidence that the primary function is communicative. *Human Communication Research*, *14*(3), 275–299.

Bavelas, J. B., Black, A., Lemery, C. R. & Mullett, J. (1986). "I show how you feel": motor mimicry as a communicative act. *Journal of Personality and Social Psychology*, *50*(2), 322–329.

Bavelas, J. B., Coates, L. & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, *79*(6), 941–952.

Bavelas, J. B., Coates, L. & Johnson, T. (2002). Listener responses as a collaborative process: the role of gaze. *Journal of Communication*, *52*(3), 566–580.

Bavelas, J. B., Gerwing, J., Healing, S. & Tomori, C. (2016). Microanalysis of face-to-face dialogue. An inductive approach. In C. A. VanLear & D. J. Canary (Eds.), *Researching interactive communication behavior: a sourcebook of methods and measures* (pp. 129–157). Thousand Oaks, CA: Sage.

Bavelas, J. B. & Healing, S. (2013). Reconciling the effects of mutual visibility on gesturing: a review. *Gesture*, *13*(1), 63–92.

Benjamin, W. (1936). *The work of art in the age of its technological reproducibility*. Penguin, 2008.

Bennett, S. (1997). *Theatre audiences: a theory of production and reception* (2nd ed.). London: Routledge.

Broth, M. (2011). The theatre performance as interaction between actors and their audience. *Nottingham French Studies*, *50*(2), 113–133.

Brown-Schmidt, S. & Tanenhaus, M. K. (2008, June). Real-time investigation of referential domains in unscripted conversation: a targeted language game approach. *Cognitive Science: A Multidisciplinary Journal*, *32*(4), 643–684.

Bull, P. & Noordhuizen, M. (2000). The mistiming of applause in political speeches. *Journal of Language and Social Psychology*, *19*(3), 275–294.

Bull, P. & Wells, P. (2002). By invitation only? An analysis of invited and uninvited applause. *Journal of Language and Social Psychology*, *21*(3), 230–244.

Carlson, M. A. (2003). The Theatre Journal auto archive: Marvin Carlson. *Theatre Journal*, *55*(1), 207–211.

Carù, A. & Cova, B. (2006). How to facilitate immersion in a consumption experience: appropriation operations and service elements. *Journal of Consumer Behaviour*, *5*(1), 4–14.

Causey, M. (1999). The screen test of the double: the uncanny performer in the space of technology. *Theatre Journal*, *51*(4), 383–394.

Chortle. (2015). *Chortle Awards – Winners and Nominees* (Online).
Available from: `http://www.chortle.co.uk/awards/2014/01/10/19382/chortle_awards`.

Chovil, N. (1991). Social determinants of facial displays. *Journal of Nonverbal Behavior*, *15*(3), 141–154.

Clark, H. H. (1996, May). *Using language*. Cambridge University Press.

Clark, H. H. & Schaefer, E. F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes*, *2*(1), 19–41.

Clark, H. H. & Wilkes-Gibbs, D. (1986, February). Referring as a collaborative process. *Cognition*, *22*(1), 1–39.

Clark, T., Lisboa, T. & Williamon, A. (2014, February). An investigation into musicians' thoughts and perceptions during performance. *Research Studies in Music Education*, *36*(1).

Clayman, S. E. (1993). Booing: the anatomy of a disaffiliative response. *American Sociological Review*, *58*(1), 110–130.

Couldry, N. (2004, October). Liveness, "reality," and the mediated habitus from television to the mobile phone. *The Communication Review*, *7*(4), 353–361.

Deller, R. (2011, May). Twittering on: audience research and participation using Twitter. *Participations: Journal of Audience & Reception Studies*, *8*(1).

Dixon, S. (2007, February). *Digital performance: a history of new media in theater, dance, performance art, and installation*. Cambridge, Mass.: The MIT Press.

Dmochowski, J. P., Bezdek, M. A., Abelson, B. P., Johnson, J. S., Schumacher, E. H. & Parra, L. C. (2014, July). Audience preferences are predicted by temporal reliability of neural processing. *Nature Communications*, *5*, 379.

Dolan, J. (2006, June). Utopia in performance. *Theatre Research International*, *31*(02), 163.

Double, O. (2005). *Getting the joke: the inner workings of stand-up comedy*. London: Methuen.

Ernst, A., Ruf, T. & Küblbeck, C. (2009). A modular framework to detect and analyze faces for audience measurement systems. In *GI Jahrestagung* (pp. 3941–3953). GI.

Fernández-Dols, J.-M. & Ruiz-Belda, M.-A. (1995). Are smiles a sign of happiness? Gold medal winners at the Olympic Games. *Journal of Personality and Social Psychology*, *69*(6), 1113–1119.

Fischer-Lichte, E. (2008). *The transformative power of performance: a new aesthetics*. London: Routledge.

Fischer-Lichte, E. (2010). Culture as performance - developing a concept of performance. In L. Sætre, A. M. Gullestad & P. Lombardo (Eds.), *Exploring textual action* (pp. 123–140). Aarhus University Press.

Fischer-Lichte, E. (2014). *The Routledge introduction to theatre and performance studies* (M. Arjomand & R. Mosse, Eds.). Abingdon, Oxon: Routledge.

Freshwater, H. (2009). *Theatre and audience*. London: Palgrave Macmillan.

Fridlund, A. J. (1991). Sociality of solitary smiling: potentiation by an implicit audience. *Journal of Personality and Social Psychology*, *60*(2), 229–240.

Frischen, A., Bayliss, A. P. & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694–724.

Gardair, C. (2013). *Assembling audiences*. Queen Mary University of London: Queen Mary University of London.

Gerwing, J. (2008). *Quantifying infant social responsiveness: microanalysis of home videos of a set of triplets for early indications of autism* (Doctoral dissertation, University of Victoria).

Gilbert, N. (2007). Computational Social Science: agent-based social simulation. In D. Phan & F. Amblard (Eds.), *Agent-based modelling and simulation in the social and human sciences* (pp. 115–134). Oxford: The Bardwell Press.

Goffman, E. (1959). *The presentation of self in everyday life*. New York: Anchor Books.

Goffman, E. (1964). The neglected situation. *American Anthropologist*, *66*(6), 133–136.

Goffman, E. (1981). Footing. In *Forms of talk* (pp. 124–159).

Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language: studies in ethnomethodology* (pp. 97–121). New York: Irvington Publishers.

Goodwin, C. (1981). *Conversational organization: interaction between speakers and hearers*. London: Academic Press.

Goodwin, C. (1986). Audience diversity, participation and interpretation. *Text - Interdisciplinary Journal for the Study of Discourse*, *6*(3).

Hagedorn, J., Hailpern, J. & Karahalios, K. G. (2008, May). VCode and VData: illustrating a new framework for supporting the video annotation workflow. In *Proceedings of the working conference on Advanced Visual Interfaces* (pp. 317–321). New York: ACM.

Harrison, S. & Dourish, P. (1996). Re-place-ing space: the roles of place and space in collaborative systems. In *Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work* (pp. 67–76). New York: ACM.

Healey, P. G. T., Frauenberger, C., Oxley, R., Schober, M. F. & Welton, M. (2009). Engaging audiences. In *CHI 2009 Workshop: Crowd Computer Interaction*.

Heath, C. & Luff, P. (2010). Orders of bidding: organising participation in auction of fine art and antiques. In N. Llewellyn & J. Hindmarsh (Eds.), *Organization, interaction and practice: studies of real-time work and organizing* (pp. 119–139). Cambridge University Press.

Heath, C. & vom Lehn, D. (2004, December). Configuring reception: (dis-)regarding the 'spectator' in museums and galleries. *Theory Culture and Society*, *21*(6), 43–65.

Heath, C., vom Lehn, D. & Osborne, J. (2005). Interaction and interactives: collaboration and participation with computer-based exhibits. *Public Understanding of Science*, *14*(1), 91–101.

Heritage, J. (1984). *Garfinkel and ethnomethodology*. Cambridge: Polity Press.

Heritage, J. (2005). Conversation analysis and institutional talk. *Handbook of language and social interaction*, 103–147.

Heritage, J. & Greatbatch, D. (1986). Generating applause: a study of rhetoric and response at party political conferences. *American Journal of Sociology*, *92*(1), 110–157.

Highfield, T., Harrington, S. & Bruns, A. (2013). Twitter as a technology for audiencing and fandom: the #Eurovision phenomenon. *Information, Communication & Society*, *16*(3), 315–339.

Hymes, D. (1974). *Foundations in sociolinguistics. an ethnographic approach*. Philadelphia: University of Pennsylvania Press.

Jokinen, K., Nishida, M. & Yamamoto, S. (2010). On eye-gaze and turn-taking. In *Proceedings of the 2010 Workshop on Eye Gaze in Intelligent Human Machine Interaction* (pp. 118–123). New York: ACM.

Katevas, K., Healey, P. G. T. & Harris, M. T. (2014, November). Robot stand-up: engineering a comic performance. In *Proceedings of the Workshop on Humanoid Robots and Creativity at the 2014 IEEE-RAS International Conference on Humanoid Robots*.

Katevas, K., Healey, P. G. T. & Harris, M. T. (2015). Robot comedy lab: experimenting with the social dynamics of live performance. *Frontiers in Psychology*, *6*(1253), 322.

Kendon, A. (1990). Spatial organization in social encounters: the F-formation system. In *Conducting interaction: patterns of behaviour in focused encounters*. Cambridge University Press.

Kohler, K. J. (2008, January). 'Speech-smile', 'speech-laugh', 'laughter' and their sequencing in dialogic interaction. *Phonetica*, *65*(1-2), 1–18.

Kraut, R. E. & Johnston, R. E. (1979). Social and emotional messages of smiling: an ethological approach. *Journal of Personality and Social Psychology*, *37*(9), 1539–1553.

Küblbeck, C. & Ernst, A. (2006, June). Face detection and tracking in video sequences using the modified census transformation. *Image and Vision Computing*, *24*(6).

Kuhlen, A. K. & Brennan, S. E. (2013, February). Language in dialogue: when confederates might be hazardous to your data. *Psychonomic bulletin & review*, *20*(1), 54–72.

Lee, Y.-A. (2007). Third turn position in teacher talk: contingency and the work of teaching. *Journal of Pragmatics*, *39*(6), 1204–1230.

Livingston, E. (1987). *Making sense of ethnomethodology*. Routledge & Kegan Paul.

Lockyer, S. & Myers, L. (2011, November). 'It's about expecting the unexpected': live stand-up comedy from the audiences' perspective. *Participations: Journal of Audience & Reception Studies*.

Loomis, J. M., Kelly, J. W., Pusch, M., Bailenson, J. N. & Beall, A. C. (2008). Psychophysics of perceiving eye-gaze and head direction with peripheral vision: implications for the dynamics of eye-gaze behavior. *Perception*, *37*(9), 1443–1457.

MacDorman, K. F. & Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies*, *7*(3), 297–337.

Mackintosh, I. (1993). *Architecture, actor and audience*. Routledge.

Mann, R. P., Faria, J., Sumpter, D. J. T. & Krause, J. (2013, August). The dynamics of audience applause. *Journal of The Royal Society Interface*, *10*(85), 20130466–20130466.

Mutlu, B., Forlizzi, J. & Hodgins, J. (2006, December). A storytelling robot: modeling and evaluation of human-like gaze behavior. In *6th IEEE-RAS International Conference on Humanoid Robots* (pp. 518–523).

Palmer, J. (1994). *Taking humour seriously*. London: Routledge.

Pasquier, D. (2015, May). 'The cacophony of failure': being an audience in a traditional theatre. *Participations: Journal of Audience & Reception Studies*, *12*(1).

Pennebaker, J. W. (2010, June). Perceptual and environmental determinants of coughing. *Basic and Applied Social Psychology*, *1*(1), 83–91.

Pertaub, D.-P., Slater, M. & Barker, C. (2002, February). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators and Virtual Environments*, *11*(1).

Phelan, P. (1993). The ontology of performance: representation without reproduction. In *Unmarked: the politics of performance* (pp. 146–166). London: Routledge.

Pitts, S. E., Dobson, M. C., Gee, K. & Spencer, C. P. (2013, November). Views of an audience: understanding the orchestral concert experience from player and listener perspectives. *Participations: Journal of Audience & Reception Studies*, *10*(2), 65–95.

Pomerantz, A. (1984). Agreeing and disagreeing with assessments: some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: studies in conversation analysis* (pp. 57–101). Cambridge, U.K.: Cambridge University Press.

Quirk, S. (2011, November). Containing the audience: the 'room' in stand-up comedy. *Participations: Journal of Audience & Reception Studies*, *8*(2).

Reason, M. (2004, May). Theatre audiences and perceptions of 'liveness' in performance. *Participations: Journal of Audience & Reception Studies*, *1*(2).

Reason, M. (2006). *Documentation, disappearance and the representation of live performance*. London: Palgrave Macmillan UK.

Reason, M. (2010). Asking the audience: audience research and the experience of theatre. *About Performance*, (10), 15.

Reason, M. & Sedgman, K. (2015, May). Editors' introduction: themed section on theatre audiences. *Participations: Journal of Audience & Reception Studies*, *12*(1).

Reeves, S., Sherwood, S. & Brown, B. (2010). Designing for crowds. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction*. New York: ACM Press.

Rienks, R., Poppe, R. & Heylen, D. (2010, January). Differences in head orientation behavior for speakers and listeners: an experiment in a virtual environment. *ACM Trans. Appl. Percept. 7*(1), 2:1–2:13.

Rosen, C. (2002). *Piano notes: the hidden world of the pianist*. London: Allen Lane.

Rutter, J. (1997). *Stand-up as interaction: performance and audience in comedy venues* (Doctoral dissertation, University of Salford).

Sacks, H. (1987). [1973] On the Preferences for Agreement and Contiguity in Sequences in Conversation. In G. Button & J. R. E. Lee (Eds.), *Talk and social organisation* (pp. 54–69). Clevedon: Multilingual Matters.

Scarpetta, F. & Spagnolli, A. (2009). The interactional context of humor in stand-up comedy. *Research on Language and Social Interaction*, *42*(3), 210–230.

Schegloff, E. A. (1998). Body torque. *Social Research*, *65*(3), 535–596.

Schober, M. F. & Clark, H. H. (1989, April). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*(2), 211–232.

Schoenmakers, H. (1990). The spectator in the leading role: developments in reception and audience research within theatre studies. In W. Sauter (Ed.), *New Directions in Theatre Research*. Munksgaard: Nordic theatre studies.

Scott, S. K., Lavan, N., Chen, S. & McGettigan, C. (2014, December). The social life of laughter. *Trends in Cognitive Sciences*, *18*(12), 618–620.

Silva, J. F. O. e., Soares, M. J., Miranda, L. S., Pereira, A. & Harper, N. L. (2013). Performing together?a case study of physiological stress between soloist and audience. In *Proceedings of the International Symposium on Performance Science*.

Slater, M., Pertaub, D.-P. & Steed, A. (1999). Public speaking in virtual reality: facing an audience of avatars. *Computer Graphics and Applications, IEEE*, *19*(2), 6–9.

Stevens, C. J., Dean, R. T., Vincs, K. & Schubert, E. (2014). In the heat of the moment: audience real time response to live music and dance performance. In *Coughing and clapping: investigating audience experience*. Farnham, Surrey: Ashgate.

Time Out London. (2015). *Timeout London – Angel Comedy* (Online).
Available from: `http://www.timeout.com/london/comedy/angel-comedy`.

Trueman, M. (2009, October). Is the live theatre experience dying? *The Guardian*.

Tsay, C.-J. (2013, September). Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(36), 14580–14585.

Tsui, A. B. M. (1989, December). Beyond the adjacency pair. *Language in Society*, *18*(04), 545.

Victor, Bret. (2011). *Up and down the ladder of abstraction* (Online).
Available from: `http://worrydream.com/LadderOfAbstraction/`.

Wagener, A. (2012, December). *Why do people (not) cough in concerts?* Association for Cultural Economics International.

Wang, C., Wong, J., Zhu, X., Röggla, T., Jansen, J. & Cesar, P. (2016, June). Quantifying audience experience in the wild: heuristics for developing and deploying a biosensor infrastructure in theaters. In *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*.

Wang, C., Zhu, X., Geelhoed, E., Biscoe, I., Röggla, T. & Cesar, P. (2016). How are we connected? Measuring audience galvanic skin response of connected performances. In *Proceedings of the 3rd International Conference on Physiological Computing Systems* (pp. 33–42). SCITEPRESS - Science and Technology Publications.

Webster, E. (2011, October). *Promoting live music in the UK: a behind-the-scenes ethnography* (Doctoral dissertation, University of Glasgow).

Williamon, A., Aufegger, L. & Eiholzer, H. (2014). Simulating and stimulating performance: introducing distributed simulation to enhance musical learning and performance. *Frontiers in Psychology*, *5*.

Wise, K. (2104, May). *English Touring Opera – 'Opera in Cinemas' report*. CreativeWorks London.

Woods, P. (2012). *Globe Audiences: Spectatorship and Reconstruction at Shakespeare's Globe*. Queen Mary University of London: Queen Mary University of London.

Yngve, V. H. (1970). On getting a word in edgewise. In *Papers from the Sixth Regional Meeting, Chicago Linguistics Society* (pp. 567–577). University of Chicago.

Zhang, J. R., Sherwin, J., Dmochowski, J. P., Sajda, P. & Kender, J. R. (2014). Correlating speaker gestures in political debates with audience engagement measured via EEG. In *Proceedings of the 22nd ACM International Conference on Multimedia* (pp. 387–396). New York: ACM.